FAST INTERCONNECTION ANALYSIS OF PV SYSTEMS USING VECTOR QUANTIZATION

A Dissertation
Presented to
The Academic Faculty

by

Jeremiah Deboever

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2018

FAST INTERCONNECTION ANALYSIS OF PV SYSTEMS USING VECTOR QUANTIZATION

Approved by:

Dr. Santiago Grijalva, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. A. P. Meliopoulos
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Lukas Graber
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Shabbir Ahmed
School of Industrial & Systems Engineering
Georgia Institute of Technology

Dr. Maryam Saeddifard
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: March 30th, 2018

To my parents, Evelyne and Jean-Francois,

sister and brothers, Jessica, Michael, and Nathaniel,

with all my love.

ACKNOWLEDGEMENTS

First, I would like to thank my family for supporting and encouraging me throughout my academic career. I would like to specifically thank my parents for the values they have taught me and for always encouraging me to pursue my passions that led me to this career achievement. Without their support, I wouldn't be where I am today.

I would like to thank my advisor Dr. Santiago Grijalva for his valuable guidance and mentorship throughout my PhD. He provided an enjoyable working environment by providing some freedom in my critical thinking that helped me grow as a researcher and engineer. I could not have asked for a better student-advisor relationship and I will forever be grateful. I would like to also thank my committee members for providing valuable feedback on my research: Dr. Lukas Graber, Dr. Mariam Saeedifard, Dr. A. P. Meliopoulos, and Dr. Shabbir Ahmed.

I would like to thank my current and previous labmates for their support throughout my PhD: Jouni Peppanen, Victor Chukwuka, Xiaochen Zhang, Leilei Xiong, James Thomas, Umer Qureshi, Sadegh Vejdan, and David Diaz. Specifically, Xiaochen and Umer provided me with valuable feedback in the development of the ideas presented in this dissertation.

This research was partially supported by the DOE SunShot Initiative, under agreement 30691 through a collaboration with Sandia National Laboratory. I would like to thank Robert Broderick and Matthew J. Reno for their support and mentorship throughout my PhD career.

Additionally, this research was also partially supported by the Electric Power Research Institute through a direct collaboration that flourished from a summer internship. I would like to thank Jeff Smith, Jouni Peppanen, Matthew Rylander, and Lindsey Rogers for making this collaboration possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ANSI	American National Standard Institute
CAP	Capacitor bank
EMT	Electro-Magnetic Transient
EV	Electric Vehicle
ESS	Energy Storage System
FDPF	Fast Decoupled Power Flow
IEEE	International Electrical and Electronics Engineers
LTC	Load Tap Changer
PV	Photovoltaic
QSTS	Quasi-Static Time-Series
SUB	Substation LTC
REG	Voltage Regulator
RMS	Root Mean Square
VQ	Vector Quantization

SUMMARY

The growing interests in solar photovoltaic (PV) systems by customers and utilities has created a surge in new interconnection of distributed PV systems on the electrical grid. One of the many benefits of having solar PV installations distributed across the system is that it provides a virtually free source of energy closer to the electrical demand with zero green-house gas emission. However, solar PV systems have an intermittent source of energy due to the diurnal and variable nature of solar irradiance that can challenge the traditional operation of distribution system. Furthermore, as the solar PV penetration level becomes significant on a distribution feeder, some problems may arise due to the reverse power flow and the high frequency variability they introduce. Thus, it is important for system planners to conduct impact studies prior to the connection of systems of significant size to avoid any unforeseeable operational issues which could reduce the lifespan or even permanently damage equipment.

Traditionally, impact studies for new interconnection of solar PV systems have been scenario-based by simulating steady-state power flow of snapshots in time, such as peak load period, which historically has been the extreme system conditions. However, scenario-based impact studies may not be adequate to accurately analyze the interactions of new interconnected resources since extreme conditions may occur at times of lower demand. Moreover, such simulation technique is only as comprehensive as the set of scenarios and can only estimate static impacts on the system. On the other hand, quasi-static time-series (QSTS) simulation introduces a temporal dimension to the simulation by chronologically solving static power flows over a time horizon. As a result, this simulation

technique can estimate time-dependent impacts such as the operation of voltage-regulating equipment or the duration/frequency of any extreme conditions experienced on the feeder. A yearlong QSTS simulation can capture the seasonal variation of the demand and the solar PV system while a 1-second resolution can accurately model the short delays in controllers. Although this simulation technique provides a much more comprehensive impact analysis for new solar PV interconnections, its computational burden and data requirements prohibit it from a practical use by the industry. Computing a yearlong QSTS simulation at 1-second resolution on a realistic distribution feeder could take anywhere from 10 to 120 hours depending on the size of the circuit. This dissertation presents a fast time-series algorithm addressing the computational burden of QSTS simulation to make it practical for industry use.

The vector quantization algorithm presented in this dissertation can drastically reduce the computational time of QSTS simulations by leveraging the seasonal and daily behavior of the load and PV system. Instead of re-computing the unbalanced AC three-phase power flow equations with an iterative solver, the algorithm stores unique power flow solutions and reassigns them each time the very similar conditions are experienced as it progresses throughout the time horizon. A hashing function is used to rapidly access previously computed power flow solutions in the solution space. Since the hashing function is extremely quick in comparison to the iterative AC power flow solver, significant computational time reduction for QSTS simulations can be achieved. The computational time of the vector quantization algorithm is then dependent on the number of unique power flow solutions computed over the time horizon and vector quantization can further reduce the computational time by clustering similar power flow solutions together. Moreover, by

only storing computed power flows, memory requirement can also be reduced to access the data post simulation.

To make it practical for industry use, the scalability and robustness of the algorithm is demonstrated in this dissertation. The vector quantization algorithm is tested on a large distribution system feeder with 3,000 buses, multiple load and PV profiles, and multiple voltage-regulating equipment. In particular, this work demonstrates that the size of the distribution feeder does not impact the performance of the algorithm and that the algorithm can scale to model feeders with multiple load and PV profiles as well as feeders with multiple voltage-regulating devices. Finally, this work demonstrates that the vector quantization algorithm is capable of handling feeders with any type of controllable devices that would be modeled under a QSTS simulation.

This dissertation also presents an efficient method to quantize the power injection profiles to optimize the speed and accuracy of the vector quantization algorithm. Through a voltage-sensitivity analysis, the profiles are quantized based on their impact on the different controllers on the feeder. Simulation results shows that the method is on the pareto front maximizing the accuracy in terms of the computational time. Since this sensitivity-based method is circuit-based, the vector quantization algorithm is standalone and can be directly implemented in commercial-grade distribution planning software.

The key contribution of this dissertation is the formulation of an algorithm capable of: (i) drastically reducing the computational time of QSTS simulations, (ii) accurately modeling distribution system voltage-control elements with delays and deadbands, and (iii) efficiently compressing result time series data for post-simulation analysis. Significant

efforts are dedicated towards the scalability and robustness of the algorithm with the objective to implement it in commercial distribution planning software such as CYME or OpenDSS. Finally, a sensitivity-based analysis makes the algorithm circuit-specific for it to be standalone and accessible to most users. The developed vector quantization algorithm makes QSTS analysis a practical tool for utility planners and operators.

CHAPTER 1. INTRODUCTION

1.1 Electrical Grid of the Future

The electrical grid in the United States has been rapidly evolving over the past decade due to the growing interest in renewable energy resources. The overall consumption of energy from hydroelectric, solar, wind, biofuels, geothermal, etc. has seen an increase of 55% between 2006 and 2016 [1]. This growth corresponds to an array of factors such as public awareness, policy changes, cost reduction, electricity price increase and market acceptance of emerging technologies. These renewable energy resources have been integrated at various levels of the electrical grid based on their sizes and the benefits they would provide to the system. Large systems are often integrated at the transmission (here >100kV) and sub-transmission (here >33kV) level where the networks can host such systems to compete with existing generation to reduce the cost of electricity and green-house gas emissions. Some countries have already reached a significant penetration level of renewable energy resources. For example, more than 85% of electricity production in Germany came from renewable resources on April 30th 2017 [2]. Renewable energy resources have also been introduced at the distribution level (here <33kV) where they have the potential to provide local benefits to the system (e.g. distributed generation). Specifically, solar photovoltaic (PV) has become a very popular technology among utilities and customers because of their decreasing capital costs and available incentives [3]. As a result, the consumption of energy from solar alone saw an increase of 6,700% between 2006 and 2016 [1].

Although distributed generation provides multiple benefits, distribution system feeders with high penetration levels of distributed resources can observe reverse power flow and new distributed resources can increase the local voltage. As a result, system operators are often required to adjust the operation of conventional voltage-regulating devices to maintain voltages across the feeder within operational limits. Another important benefit of renewable energy resources is that their source of energy is unlimited for all practical purposes and that their operational costs tend to be very low. However, the challenge with certain renewable energy resources is that their source of energy is intermittent and may not be available during crucial time when electricity demand is high. More importantly, the electricity production of some renewable energy resources is highly variable which creates a significant challenge the operation of the electrical grid. For example, the generation output of solar PV systems can be highly variable and intermittent particularly during cloudy days. Again, this becomes a challenge for system operators to manage the adverse power quality and reliability impacts from solar PV and to maintain the system within their operational limits. This dissertation focuses on the impacts of new solar PV systems interconnections on distribution networks because of the impacts they can introduce and their rapid growth over the past 10 years.

The explicit impacts solar PV systems have on the electrical grid may vary from one distribution feeder to another depending on the topology and characteristics of the system as well as the control modes and set points of any voltage-regulating equipment. Thus, a detailed impact analysis is often needed prior to interconnecting any new solar PV systems of a significant size. Conventional analysis techniques focus on simulating a few power flow scenarios to study the power flow and the nodal voltage across a feeder under

certain conditions. However, the detail of such analysis is as comprehensive as the set of simulated scenarios and extreme scenarios may not be representative of realistic conditions experienced on a distribution system feeder. Because of the lack of a detailed analysis technique, some utilities have limited conservatively the number and size of new interconnections on some distribution system feeders with significant existing PV system, if there is any chance of operational issues (e.g. voltage quality). For example, Hawaii has recently experienced considerable growth in distributed solar PV system interconnections to a level that the electric utility has limited the number of new grid-tied systems due to the potential risk of negative impacts on the rest of the system [4]. Maui Electric published a map of the remaining percentage of the distributed generation capacity available on different branches of their network (Figure 1). One can observe that many sections are quickly approaching saturation not allowing customers to install new generation.

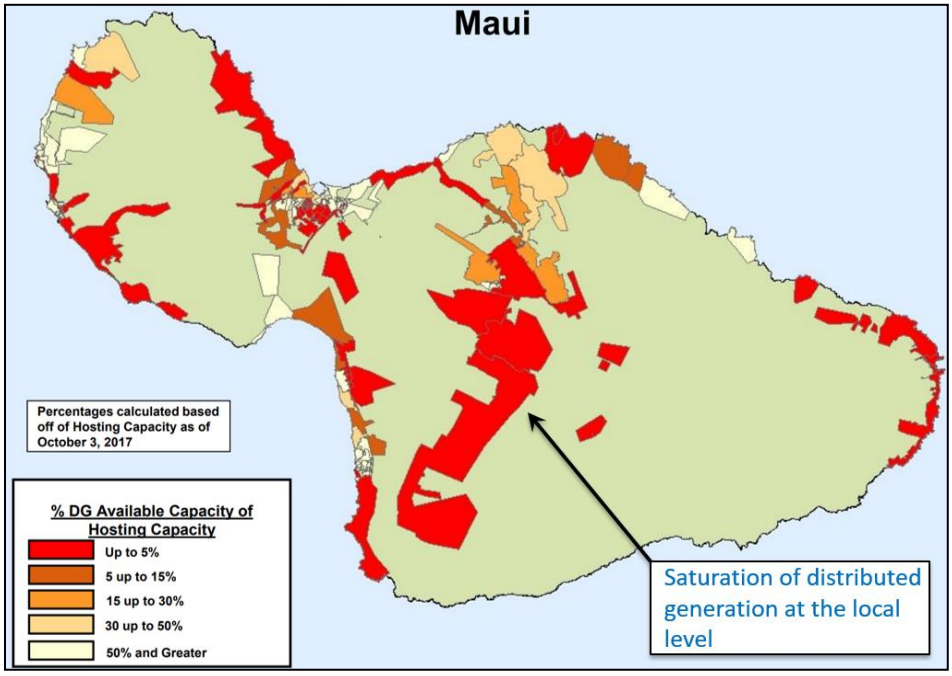


Figure 1 – Percentages of distributed generation available capacity on the island of Maui [4]

In order to achieve a higher penetration level of distributed generation on distribution feeders, system planners must study the impact that new interconnections would have on the operation of a distribution feeder. Without a detailed simulation technique, the impacts created by a new interconnection cannot be thoroughly studied and could result in unexpected operational issues. In the following section, conventional interconnection studies used by utilities are discussed in detail and a new simulation technique suggested by IEEE is highlighted.

1.2 Simulation Techniques for new PV interconnections on distribution feeders

1.2.1 Conventional interconnection studies

Distribution networks with high penetration levels of solar PV can experience voltage quality issues or thermal violations that can reduce the lifespan of some equipment. Thus, impact analyses for new interconnections are often required for systems of significant size or feeders with existing distributed energy resources. Conventionally, solar PV interconnection studies have focused on steady-state power flow simulations, harmonic analysis, and system protection studies [5]. These types of studies have traditionally been sufficient for distribution system planners to design feeder layouts, plan expansions, consider upgrades, and determine distribution system control settings. The driving assumption for these studies was a slowly-varying, fairly predictable load, which enabled effective design by considering only static, worst-case scenarios. However, solar PV systems and associated novel capabilities such as advanced inverters do not match this assumption, requiring new analysis methods. Commercial distribution planning software tools provide the capability to perform steady-state power flow of snapshots in time, such

as peak load period, which historically has been the extreme system conditions. Traditional snapshot tools and methods may not be adequate to accurately analyze the interactions of new interconnected resources. For instance, a distribution system with solar PV may experience over-voltages at times of lower demand or may be affected the operation of switchable voltage-regulating devices, such as load tap changers (LTCs) or capacitor banks. As more resources are introduced to the system, the conventional steady-state scenario-based power flow modeling is also not sufficient for studying the impact of distributed solar PV because it cannot capture time-dependent effects of controllers with delays and deadbands. Specifically, the operation voltage regulators and capacitor banks with local controls may be significantly affected by solar PV power output variations. This can lead to rapid voltage changes on the circuit, and correspondingly excessive number of actions of the controller, resulting in shorter life of the asset. On the other hand, chronological time-series simulation can capture the possible conditions as well as the detailed effects of controllers, and hence tools with this capability are needed to provide a comprehensive understanding of the impact of PV interconnections on distribution feeders.

1.2.2 Quasi-Static Time-Series (QSTS) simulation

A recent IEEE guide (P1547.7 D110) on *'conducting distribution impact studies for distributed resource interconnection'* discusses four types of special system impact studies [5]: (1) dynamic simulation, (2) electromagnetic transient (EMT) simulation, (3) harmonic and flicker study, and (4) quasi-static simulation. Dynamic simulations model the system dynamics to understand stability issues or voltage and frequency ride-through capabilities [5]. EMT simulations study the impact that the interconnection of a new distributed resource has on the protection design and fault analysis [5]. Harmonic and

flicker studies provide insights on the feeder's power quality [5]. Last, quasi-static simulations model the operation of controllable devices on a feeder, the power protection coordination, and the voltage regulation and reactive power management [5]. This dissertation focuses on quasi-static time-series (QSTS) simulations to study the impact of new PV interconnections. QSTS analysis is used to analyze the feeder impact of solar PV and load variations over a given time-period. In particular, QSTS simulation is used to capture the interdependencies between simulation time steps that scenario-based snapshot power flow simulation cannot accurately capture. QSTS simulation is best defined by the IEEE draft guide:

> *"Quasi-static simulation refers to a sequence of steady-state power flow, conducted at a time step of no less than 1 second but that can use a time step of up to one hour. Discrete controls, such as capacitor switch controllers, transformer tap changers, automatic switches, and relays, may change their state from one step to the next. However, there is no numerical integration of differential equations between time steps."* [5]

QSTS simulation provides a more detailed analysis of the impact of new PV interconnections and offers many practical advantages and uses over conventional tools:

1) It is not limited to specific time periods, such as peak load time, which may no longer be the most critical times with high penetrations of solar PV, and models all realistic power flow scenarios based on historical data.

2) It can study the operation of discrete controller algorithms, such as voltage-regulating devices or PV advanced inverter (e.g. volt-var control), and the interactions between the different controllable equipment.

3) It can simulate impacts of fast fluctuations that are caused by highly variable resources such as distributed solar PV.

4) It can study the interaction between the daily changes in load and PV output and perform energy and loss evaluations over actual profiles of load and generation.

5) It can determine temporary over-voltage conditions that can occur before voltage-regulating device operations.

6) It can calculate the time duration of extreme conditions, such as the number of hours a customer is expected to see an over-voltage condition or the amount of time a conductor or transformer is overloaded each year.

Many potential impacts created with new PV interconnections, such as the duration of voltage violations or the increase in voltage-regulating device operations, cannot be accurately evaluated without time-series simulation. Furthermore, snapshot study methods which only analyze peak periods or a peak variability day often lead to over-estimation of normal operating with issues. Paired with accurate load and generation time-series data or models, a quasi-static time-series analysis (QSTS) can realistically quantify not only the magnitude but also the frequency and duration of an impact [1]. A comparison of the type of study that each simulation technique is capable of conducting is summarized in TABLE 1.

TABLE 1. Comparison of the capability of conventional snapshot simulation and QSTS simulation

Impact study type	Conventional snapshot simulation	Quasi-Static Time-Series simulation
Power flow studies		
Thermal violations	Good	Great
Thermal violation duration	N/A	Great
Reactive power management	Good	Great
Power losses	N/A	Great
Voltage quality		
Voltage extremes (highest/lowest)	Good	Great
Voltage violation duration	N/A	Great
Voltage regulating devices		
Controller operation (# of actions)	Poor	Great
Interactions between controllers	N/A	Great
Smart devices operation		
Smart inverters (volt/var, volt/watt, etc.)	Good	Great

Without a temporal dimension, the conventional snapshot scenario-based simulation can only adequately approximate time-independent impacts but cannot study any impacts that are time-dependent (e.g. voltage violation duration). Furthermore, even though some impacts can be approximated with snapshot scenario-based simulations [6], it may not be possible to simulate all realistic scenarios that a feeder will experience. The advantage of the QSTS simulation is that it introduces a temporal dimension to the simulation that allows it, for example, to realistically model the different voltage-regulating equipment with their respective controller logic. Hence, QSTS simulation is a critical simulation technique for comprehensive interconnection studies, as recommended by IEEE [5].

A major challenge to using QSTS simulation is the computational burden associated with it. A yearlong simulation is recommended to model both the seasonal

variation of the load and the high-frequency fluctuations of solar PV while a 1-second resolution is recommended to model controllers with short delays [7]. However, this represents running 31.5 million chronological power flows for 10 to 120 hours on conventional computers, depending on the complexity and size of the circuit, which is far too slow for PV interconnection impact studies. As more historical data becomes available, the computational burden associated with running QSTS simulation is one of the reasons why it has not been widely used by the industry. Reducing the computational time of the simulation will provide a new tool for system planners to perform accurate PV interconnection studies, allowing higher penetrations of solar PV systems to be connected on the distribution system.

1.3 Research Objectives

The objective of this dissertation is to address the increasing need for a fast and comprehensive interconnection analysis technique (QSTS simulation) for the integration of solar PV systems on distribution system networks. The research in this dissertation overcomes the limitation of conventional scenario-based simulations, which provide only a snapshot perspective on the impact of solar PV and which cannot capture any temporal impacts such as controller actions or duration of extreme conditions.

The challenge in the research presented in this dissertation is that while quasi-static time-series simulation provides a more comprehensive impact analysis, the computational burden of such simulation prohibits its practical use by the industry. Hence, this research studies the challenges to reduce the computational time of QSTS simulations and presents a fast time-series algorithm using vector quantization. The research focuses on the impacts

of centralized and distributed solar PV systems on distribution networks and more specifically on the operation of voltage-regulating equipment on distribution feeders. The developed simulation technique is also scalable and robust to real utility distribution feeder model complexities including model size, the number and types of controllable elements, or the number of PV systems for the technique to be implementable within existing distribution planning software and used by distribution system planners and operators.

1.4 Outline of Chapters

A literature survey and the research context of QSTS simulations are presented in Chapter 2. First, the motivation of QSTS simulations is studied through a literature survey in terms of the type of distributed energy resource and the type of system impacts QSTS simulation can analyze. Second, six challenges in reducing the computational burden of QSTS simulations are discussed in detail. Last, a survey of existing fast QSTS simulation algorithms is presented to highlight the gap in the literature and place this research into context.

Chapter 3 proposes a fast QSTS simulation algorithm capable of reducing the computational burden while modeling the various complexities of distribution feeders. The chapter first discusses how QSTS simulations are traditionally performed for distribution feeders with voltage-regulating devices. A fast QSTS simulation algorithm is then proposed leveraging similarities in power flow solutions to reduce the computational time of the simulation using vector quantization and the algorithm is tested on a modified IEEE 13 bus circuit.

Chapter 4 discusses the scalability of the proposed fast QSTS simulation algorithm. The concerns regarding of the size of the feeder, the number of load/PV profile, the number of controllable elements, and the type of controllers are addressed and the algorithm is adapted to model most distribution feeders. The scalability of the algorithm is demonstrated on a large realistic distribution feeder (~3,000 buses) as well as on a feeder with smart inverters.

Chapter 5 discusses the effect of vector quantization on the accuracy of the QSTS simulation. Two different vector quantization strategies are presented and a voltage-sensitivity strategy is proposed to optimize computational speed and simulation accuracy. The performance of all three strategies are compared on the IEEE 13-bus case and a large distribution feeder test case.

Chapter 6 summarizes the research presented in this dissertation and highlights the contributions and significance of this research. Furthermore, this chapter also discusses potential future work directions that this research has created. A list of the publications that resulted from this dissertation can be found in Appendix A.

CHAPTER 2. LITERATURE SURVEY AND RESEARCH CONTEXT

2.1 Motivation for Quasi-static time-series simulation

Time-series power flow simulations has been discussed in the literature for impact studies of different resources: solar PV [6], [8]–[19], wind [20], [21], electrical vehicles (EV) [22], [23], and energy storage systems (ESS) [24]. QSTS simulation is also used for impact studies of control schemes in different power equipment: smart inverters [12], [17], [18], [25], and voltage regulating devices [15], [26]–[29]. The common objective for using QSTS analysis in all these studies is to capture the time-dependent impacts of various resources on distribution feeder. QSTS simulation can be used to perform various types of studies on a feeder, such as studying the impact of solar PV systems or control schemes on the voltage quality [6], [8]–[17], [23], [25]–[27]. The sequential time-series simulation can determine the range of voltage magnitude as well as the duration of any voltage violations on the circuit. QSTS simulation is also used to study the operation of voltage regulating devices [6], [8]–[10], [13], [15], [16], [18], [23] caused by large power flow fluctuation that certain resources create. Other types of studies performed with QSTS analysis includes: equipment loading assessment [20], [24], system losses [9], [16], [19], [22], [24], [26], [28], or power flow direction [6].

The time-step resolution and time horizon of the QSTS analysis varies based on the type of study performed and the resource studied. The need of high resolution (seconds) to study the impact of PV systems based on its high variability nature is described in [13]. On

the other hand, wind, EVs, and ESS impact studies have a minutes-to-hour resolution due to the slower variation in power injection of wind farms and periodic charging schedule of storage. In addition, the time resolution of the QSTS simulation should be below the fastest delay in any devices with discrete controls on the feeder to ensure accurate representation of the device's operation [14]. The type of study is also a factor in the time-step resolution used for QSTS. Reference [16] discusses a general recommendation of hourly resolution for energy impact analyses, minutes for steady-state overvoltage studies, and seconds to minutes for voltage fluctuations. In [7], a yearlong simulation with a 1-second resolution is recommended to accurately capture the operation of controllers with delays although acceptable error can be achieved with resolution up to 5 seconds. Therefore, a comprehensive impact study requires a yearlong simulation at a 1-second granularity [7] to represent both the seasonal variation of the load and the high-frequency fluctuations of distributed PV. This represents running 31.5 million chronological power flows for 10 to 120 hours on conventional computers, depending on the complexity and size of the circuit, which is far too slow for PV interconnection impact studies. The computational burden associated with running QSTS simulation is one of the reasons why it has not been widely used by the industry. Reducing the computational time of the simulation will provide a new tool for system planners to perform accurate PV interconnection studies, allowing high penetrations of solar energy to be connected on the distribution system.

2.2 Challenges for fast QSTS simulations

There are several challenges in reducing the computational time of QSTS simulations due to the discontinuous and nonlinear nature of the simulation [30]. Six of the most

significant challenges are briefly discussed in the following sub-sections and a more detailed discussion can be found in [30].

2.2.1 Challenge 1: Number of power flow to solve

Challenge statement – A yearlong QSTS simulation at 1-second granularity sequentially solves 31.5 million static power flows, which requires significant computational time.

Iterative power flow solvers have been researched and implemented since the 1950s [31], [32]. They are at the core of numerous power system analyses; consequently, significant effort has been devoted to improving their simulation speed (i.e. reduce the number of iterations or accelerate each iteration) since then. Fast iterative algorithms for distribution systems are already implemented in commercial software such as CYME and in open source packages such as OpenDSS, or GridLAB-D. Significant computational time reductions of the iterative solver can be theoretically achieved by reducing the number of iterations, but progress in this direction has already been pursued and solvers have already been optimized to converge with a small number of iterations within the scope of QSTS simulations. This was tested in OpenDSS with a fast-varying PV profile and a slow-varying load profile. Each time-step of a yearlong QSTS simulation at 1-second granularity converged in 2.0080 iterations on average: one iteration to compute the solution and one to check if it converged [33]. Not only do current solvers quickly converge, the initial iteration uses the solution from the previous time step in QSTS simulations, and since the variation is minimal from time-step to time-step, the algorithm converges quickly. Thus, significant time reduction cannot be achieved from reducing the number of iterations.

On the other hand, one alternative to reducing the computational time is to apply approximations to the power flow equations or to their solution methods. Some approximation methods consist of entirely linearizing the set of equations, while others find ways to approximate the Jacobian matrix to decrease the CPU time spent on matrix factorization. Numerous power flow approximation methods are presented in the literature including but not limited to:

i. Dishonest Newton: Keeps the same Newton-Raphson Jacobian matrix for a given number of iterations.
ii. Fast decoupled power flow (FDPF): Assumes that line conductance and phase shifts between nodes are negligible and that voltage magnitude are close enough to not affect real power flow. As a result, the power flow equations are separated into two smaller independent systems with constant Jacobian matrices.
iii. DC power flow: Mostly used for transmission systems due to their size and properties, this fully linear approximation ignores line conductance and reactive power flows.
iv. Series impedance voltage drop approximation: Decoupled real and reactive power calculations, ignoring capacitance.
v. LinDistFlow model: Based on the minimal spanning tree optimization problem [34], this model assumes no line losses on the feeder.

However, many of these approximations do not work well with multi-phase unbalanced distribution systems (e.g. due to their low X/R line ratio), and all power flow approximations suffer in certain conditions (whether in terms of accuracy or robustness). Since standard iterative power flow solvers typically converge in two iterations in QSTS

simulations, the dishonest Jacobian method cannot provide significant computational time reduction, especially with linear equation solvers (e.g. PARDISO and KLU) splitting the symbolical and numerical factorization stages [35]. Other methods make assumptions, about either the voltage on the feeder or the reactive power flow, that introduce an error in the solution especially with reactive power flow injections (i.e. capacitor banks, VOLT/VAR inverters, etc.) or voltage regulating devices. This error can be reflected on the accuracy of various metrics reported by the QSTS simulation. Thus, the error introduced by power flow approximations poses a significant challenge in speeding up QSTS simulations.

Moreover, the interdependency between time-steps of the QSTS simulation (see Section 2.2.3) requires each time step to be solved chronologically which furthers the argument that much more important gains can be obtained by reducing the sheer number of time-steps to be solved as opposed to the CPU time of individual power flow solutions.

2.2.2 Challenge 2: Circuit complexity

Challenge statement – The set of power flow equations for an unbalanced, 3-phase system is nonlinear by nature. When considering various controller logics, the QSTS simulation becomes a discontinuous nonlinear system that can be very complex. Simplifying this system can be very challenging without having prior knowledge of how it behaves.

The nature of this discontinuous nonlinear system makes it especially challenging to predict how it will behave. For instance, the size of a PV system may or may not impact

the operation of various controllers on a feeder. Furthermore, their impact is neither continuous nor linear as shown in Figure 2.

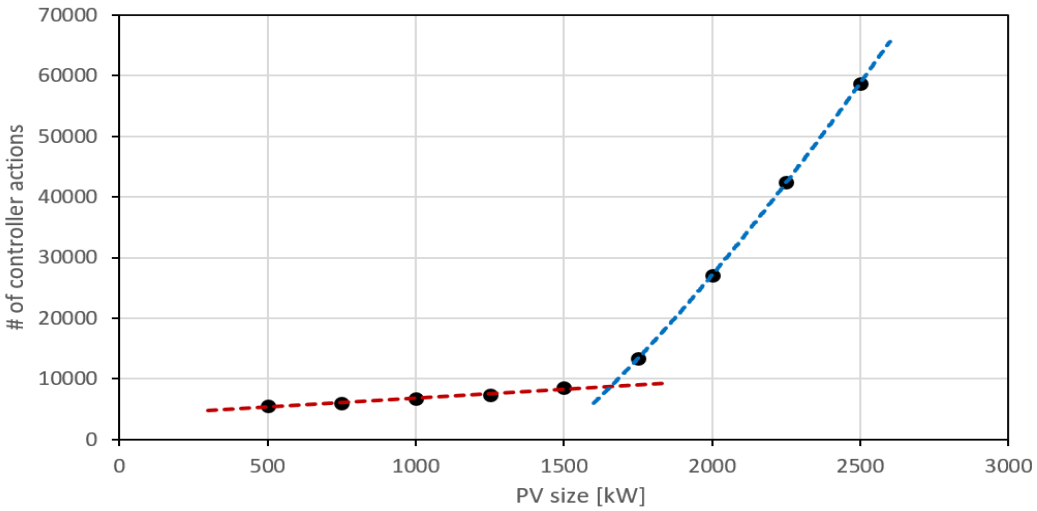


Figure 2 – Total number of controller actions on a modified IEEE 13-bus test circuit with a centralized PV system of different sizes.

As shown in the figure above, the correlation between the number of controller actions and PV size is approximately linear only until it reaches ~1500kW. This is a trend that would have been very challenging to foresee based on the characteristics of the feeder.

The size of the PV system is not the only factor that causes unpredictability in the system. Its locations, whether it is distributed or centralized, the controller settings, the location of voltage regulating devices, their interactions, etc. are a small subset of the factors impacting the operation of a distribution feeder. The circuit complexity creates a challenge of unpredictability in the system that makes modeling QSTS simulation without going through each time-step challenging.

2.2.3 Challenge 3: Time dependency between time-steps

Challenge statement – The time dependence in the controller logic of certain distribution system devices requires the QSTS simulation to be solved chronologically. For

instance, the delays and the deadbands in the controllers create a hysteresis in the state of the system. This hysteresis can be a challenge in reducing the computational time of QSTS simulations.

Controller logics in some devices on a distribution feeder can have a time dependence either by design or by nature. Delays and deadbands are often incorporated in controllers (e.g. tap changers or capacitor banks) to ignore any temporary fluctuations in power flow and avoid oscillation in their operation. Delays can filter out high frequency variations while deadbands reduce oscillations caused by their own or other devices' operation. In addition to distribution voltage-regulating devices, there can be many other devices on the distribution system with time-dependence, such as PV systems with advanced inverter controls and energy storage systems (ESS) state-of-charge (SOC) controls.

When solving each time step chronologically, the hysteresis of controllers is easily modeled through their logics. As the simulation advances second by second, the time dependence is naturally incorporated with the previous states and any delay timers. This may become a challenge for some computational time reduction approaches if the time steps are no longer solved chronologically. The controller hysteresis may not be accurately modeled or completely ignored which does not realistically represent the operation and state of the system.

2.2.4 Challenge 4: Multiple valid power flow solutions

Challenge statement – Without the historical information about previous system states, multiple valid power flow solutions exist for given power injections on a feeder due

to the deadbands and delays in controller logics. Therefore, correlating these power injections with the states of controllable devices becomes a challenge.

Deadbands are often incorporated within controllers to reduce the oscillation from their own or other devices' operation. As a result, controllable devices on a feeder can have multiple valid states within their controller limits for a given power injection (e.g. for a given demand). For instance, in voltage-regulating tap changers, system operators design the voltage deadband to include 3-5 tap positions within the thresholds to avoid oscillation. Because of these deadbands, one cannot associate a load level to a specific tap position without considering its state at a previous time step. This discontinuity in the relationship between load and system states can become a challenge in approximating controller logic models.

Because the deadbands in controllers create a hysteresis in the state of the system, approximating controller logic models can become extremely complex without modeling the actual logic of the controllers. Models based on power injections cannot be used when controllers are considered since multiple discrete system states would be valid for the same power injections. For example, a machine learning model, which seeks to establish a one to one mapping between the QSTS inputs and outputs, is not able to learn the correlation because the same inputs will yield multiple valid possible power flow solutions. This challenge can present a problem for any new QSTS algorithms that do not track the system states through time. The most intuitive solution to eliminate the effect of the multiple valid solutions is to introduce time dependence and time correlation, which itself becomes a new challenge that can be computationally cumbersome to achieve an accurate representation of the operation of the system.

2.2.5 Challenge 5: Controllable element interactions

Challenge statement – Controllable elements placed on the same phase will interact with one another. Because of their deadbands, an action in one controller caused by a small voltage approximation error in the power flow solution can create false oscillations in other controllers before it can be cleared.

Multiple voltage regulating devices can be placed on the same circuit, especially on long radial distribution feeders. Deadbands and delays in the controller of each devices are coordinated to avoid continuous oscillations between devices. However, their coordination becomes complex when PV systems introduce large fluctuations in power injections in the circuit, which can create reverse power flow. This challenge is illustrated with a modified IEEE 13-bus test circuit (see Section 3.4) with 10% and 40% PV penetration (Figure 3). Two voltage regulating devices are considered: a voltage regulating tap changer at the substation and a capacitor bank near the PV system. The regulator has 125 tap actions (instead of 3) and the capacitor banks has 52 actions (instead of 0) with the increase in PV penetration.

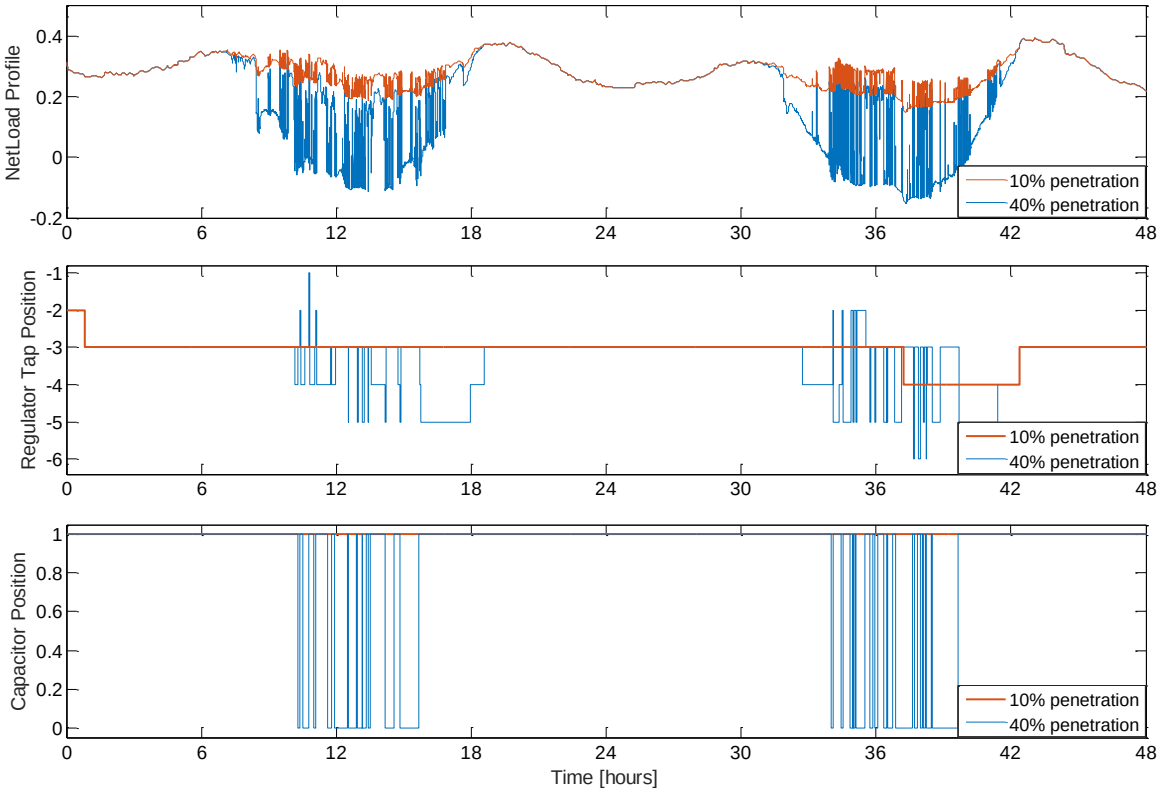


Figure 3 – Plot of the net load at the substation (normalized to peak load), regulator tap position and capacitor position for a system with 10% and 40% penetration of PV.

As expected, the tap changer will regulate the voltage to follow the daily variation of the demand. In the 10% simulation, the controller of the capacitor bank does not operate because of its delay being longer (30 sec.) than that of the tap changer (15 sec.) allowing the voltage to be regulated before the capacitor bank operates. The daily operation of the controllers is very different when a larger PV system is considered. In the 40% simulation, the capacitor bank will operate to regulate the voltage at the end of the feeder. This operation will trigger the tap changer to operate in response to the capacitor state because of the reactive power injection variation. Since the capacitor bank is more sensitive to the PV system, its operation will increase and consequently increase the operation of the tap changer.

Modeling the interactions between the operations of these voltage regulating devices can be difficult to predict especially when they interact with one another. The deadbands and delays in the controller logics are designed to avoid oscillation under specific conditions. However, new interconnections can disrupt this balance and create emergent behaviors with considerable impacts on the feeder.

Another aspect of the challenge with controller interactions is cascading errors. Because of the deadbands in the controllers, the controllable element may trigger a change and remain in that state for an extended period of time. As a result, the operation of other controllers can be significantly impacted by it. For example, the state of a capacitor bank, which is a reactive power injecting device, can impact the operation of an upstream tap changer. Because of the multiple valid power flow solutions discussed in Section 2.2.4, under the same power injection conditions, the operation of the tap changers can increase or decrease dramatically based on the state of the capacitor bank. In Figure 4, a simulation is conducted where a single regulator action is neglected and as a result triggers the capacitor to operate. This single error produces a completely different series of controller events over the following few hours causing the tap changers to record additional actions before returning to identical system states.

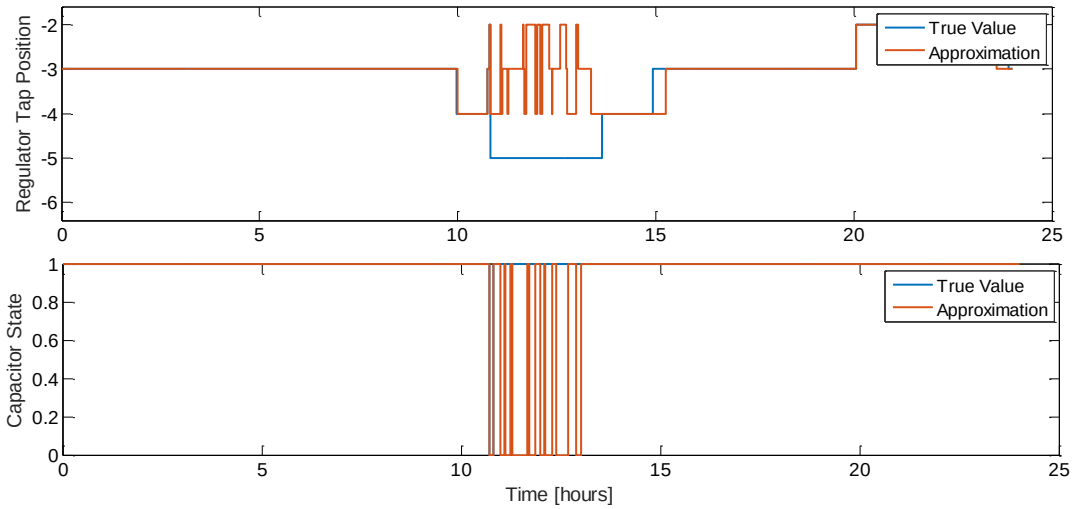


Figure 4 – States of voltage regulating devices over a 24-hour period demonstrating how the interaction between devices can create cascading errors with excess actions.

In this simulation, the cascading error only took approximately four hours to disappear but it could have easily taken a few simulation days. Speeding the QSTS simulation without going through the controller logics at each time-step is challenging because of these controller interactions. A small approximating error in the power flow solution can create a controller action that can impact how another controller will operated over a period. This is a significant challenge that affects most if not all computational time reduction approaches.

2.2.6 Challenge 6: Accurate analysis for extended time-horizon simulation

Challenge statement – In order to characterize the impact of a new resource on a feeder, various metrics (e.g. number of tap actions or voltage violations) can be computed a-posteriori based on the time-series solutions. However, a large amount of data is often required to fully understand its impact. For instance, monitoring voltage violations would require recording voltage quantities for all the nodes in the system for all the time points (e.g. 31 million time-points for 10k unbalanced nodes). In addition, the accuracy of each

reported metric may be impacted differently based on the approach taken to reduce the computational burden.

The amount of data to be recorded is dependent on the objective behind the QSTS simulation, whether it is to study the impacts on voltage quality or the operation of controllable devices. The analysis from a QSTS simulation can be categorized into two types of data: discrete metrics or time series measurements.

Discrete metrics, such as number of controller actions or total power losses, can be recorded as aggregating values at each time-steps or later computed by recording time-series data to process post simulation. Obviously, there is an advantage for both approaches. Only recording aggregated values does not have a significant memory requirement but will not allow further analysis besides the final discrete metric. On the other hand, recording time-series data requires significant memory but allows post simulation analysis.

Recording data at each time-step can increase the computational time either because of the sheer amount of data (i.e. voltage magnitude) or because of a necessary logic (i.e. tap change if-statement). More specifically, recording time-series voltage measurements for large distribution feeders (500+ buses) may not be possible without running the simulation in sequences. When the power flow solver is contained in its own dynamic-link library (DLL), non-negligible computational overhead may also result from the transfer of large amount of data at each time-step between the solver and the main application.

The purpose behind running a QSTS simulation is to understand the operation of a distribution feeder under certain conditions. However, it may be challenging to provide a

clear understanding of the system without having multiple metrics and/or time-series measurements to analyze. Data management can become a challenge based on the approach taken to reduce the computational time.

2.3 Review of fast QSTS methods

A one-year simulation at one second granularity represents 31.5 million chronological power flows. The computational burden of this simulation limits the capability to rapidly simulate multiple PV sizes, locations or configurations onto a realistic feeder. Computational time reduction of QSTS simulations can be achieved by either improving the speed of each power flow, or by reducing the number of power flows solved. A handful of publications specifically discussed shortening the QSTS analysis fall within one of the two approaches [9], [36]–[38]. Since the speed of the iterative power flow solver is proportional to the number of buses in a feeder [30], efforts have focused on simplifying the feeder to a small number of buses or dividing the feeder into sub-networks to solve them in parallel. The simplification of distribution feeders to improve the speed of PV impact studies was demonstrated in [38], [39] by reducing a 1262-bus feeder into a 13-bus circuit with highly accurate results for the preserved buses of interest. A-Diakoptics methodology for multicore power flow simulation was presented in [36] to divide large distribution grids into subnetworks, each being solved in parallel to reduce computational time with minimal error.

While the methods discussed above investigated the computational time of QSTS simulations by improving the speed of each power flow, the method presented in this dissertation focuses on the time-series algorithm aspects of QSTS and does not change the

speed of an individual power flow. Reference [7] investigated reducing the resolution or length of the simulation and an ~80% computational time reduction was achieved. Although time reduction is limited, simulation results showed that reducing the resolution provided better time reduction than reducing the time horizon of the simulation. Two other publications have investigated shortening QSTS simulations by reducing the number of computed power flows [9], [37]. Non-uniform vector quantization of load profiles, PV profiles and slack voltage profile is investigated in [37] to shorten time-series power flow simulation with time savings between 50%-70% and accuracy of $R^2 = 0.97$ for power losses. Similarly, clustering of load and production profiles has been proposed in [9] to reduce the number of power flow calculations. The reference discusses the integration of on-load tap-changers or storage but does not model them in the system and the voltage at the substation is modeled as a pre-determined time-series profile and not with discrete control algorithms. This omission does not allow the QSTS simulation to capture the operation of those controllable elements in response to the introduced PV system on a feeder.

While the A-Diakoptics and circuit reduction approaches reduce the computational time of solving power flows, the novel vector quantization approach proposed in this dissertation decreases the computational time by completely circumventing the need to solve many of the power flows in a QSTS simulation. One crucial aspect not discussed in the previous quantization work described above [9], [37] is the interaction of discrete-logic controllable elements that exist on most distribution systems. More specifically, voltage regulators and capacitor banks are not considered on the simulated circuit. As previously discussed, PV systems can have significant effects on the operation of those controllable elements and

can shorten their lifespan. In this dissertation, these elements and their effect on a feeder are considered and a vector quantization algorithm is proposed to solving the QSTS simulation.

CHAPTER 3. VECTOR QUANTIZATION ALGORITHM FOR FAST QSTS SIMULATION

3.1 Introduction

Although a QSTS simulation is a discrete set of chronological power flow computations representing a large number of scenarios, solving them consecutively captures time-dependent states in a circuit, such as voltage regulator time delays. At each time-step, the brute-force QSTS simulation solves the unbalanced three-phase power flow equations based on the current state of the system to determine the voltage magnitudes and angles – the power flow solution – at each node for each time-step over the time horizon. The resulting voltage and current magnitudes are then used in the controller logic of each controllable element to determine their new states. This process is repeated over the time horizon and time-series data can be recorded for analysis. Figure 5 shows a high-level flow diagram of the brute force QSTS simulation and all vectors are now defined.

Assuming that the feeder topology and other factors affecting the power flow solution (e.g. controller settings) do not change throughout the QSTS analysis, only the system demands and PV outputs affect the time-series simulation. An input vector $\boldsymbol{u}_t$ at time *t* can be defined as the combination of the vector $\boldsymbol{d}_t$ of system demands at time *t* and the vector $\boldsymbol{p}_{pv,t}$ of the power output of the PV systems on the feeder at time *t*:

$$\boldsymbol{u}_t \coloneqq [\boldsymbol{d}_t, \boldsymbol{p}_{pv,t}]. \tag{1}$$

where $\boldsymbol{d}_t \epsilon \mathbb{R}^{1\times d}$, $\boldsymbol{p}_{pv,t} \epsilon \mathbb{R}^{1\times p}$ and d and p are the number of profiles for the load and PV systems, respectively. If the topology of the feeder were to change during the simulation, an additional topology state can be added to the vector $\boldsymbol{u}_t$.

At each time-step, the algorithm captures the time-dependent state of the system by solving the unbalanced three-phase power flow equations and using the solution in the discrete logic of any controllable devices. The power flow solution output $\boldsymbol{g}_t$ can be defined by the following vector:

$$\boldsymbol{g}_t \coloneqq [\, |\boldsymbol{v}_t| \,, \boldsymbol{\theta}_t \,], \tag{2}$$

where $|\boldsymbol{v}| \epsilon \mathbb{R}^{1\times n}$ and $\boldsymbol{\theta} \epsilon \mathbb{R}^{1\times n}$ are the voltage magnitude and angle at each node and n is the number of nodes. The controller logic requires the power flow solution as an input signal to determine whether an action is needed for any controllable elements on the circuit. Because certain controllers may have a built-in delay to filter out temporary spikes, a delay accumulator is introduced in the logic to accumulate the number of consecutive time steps requiring a controller action before the action is taken. The logic determines the state of those elements $\boldsymbol{l}_t$ and the delay accumulators of each element $\boldsymbol{a}_t$ based on their previous states and the solution of the power flow $\boldsymbol{g}_t$:

$$\boldsymbol{l}_t \coloneqq [\boldsymbol{r}_t, \boldsymbol{c}_t], \tag{3}$$

$$\boldsymbol{a}_t \coloneqq \left[\, \boldsymbol{a}_{r,t}, \boldsymbol{a}_{c,t} \right], \tag{4}$$

where $\boldsymbol{r}_t \epsilon \mathbb{R}^{1\times r}$ is the discrete tap position of voltage regulators, $\boldsymbol{a}_{r,t} \epsilon \mathbb{R}^{1\times r}$ is their delay accumulator, $\boldsymbol{c}_t \epsilon \mathbb{R}^{1\times c}$ is the discrete on/off state of any capacitor bank, $\boldsymbol{a}_{c,t} \epsilon \mathbb{R}^{1\times c}$ is their delay accumulator, and r and c are the number of regulators and capacitors, respectively. Additional distribution system control elements, such as storage, can be added to the vector

in a similar fashion. The power flow solution $\boldsymbol{g}_t$ at time *t* is dependent on the current system demands, PV outputs, and the previous states of the controllable elements. Thus, a modified input vector $\boldsymbol{h}_t$ is defined as a vector of all factors that affects the unbalanced three-phase power flow solution:

$$\boldsymbol{h}_t \coloneqq [\boldsymbol{u}_t,\ \boldsymbol{l}_t]. \tag{5}$$

The reason controllable element states are considered as an input variable to the power flow is because the voltage magnitudes and angles on the feeder are dependent on the state of those voltage regulating devices. The state of the system at each time step can then be defined as the vector $\boldsymbol{x}_t$.

$$\boldsymbol{x}_t \coloneqq [\boldsymbol{h}_t, \boldsymbol{g}_t] \tag{6}$$

At the end of each time-step, this state vector can be stored partially or entirely in a time-series matrix $\boldsymbol{X}$ for post-simulation analysis. Various metrics for voltage quality, operation of voltage regulating devices or system losses can be determined based on the time-series matrix $\boldsymbol{X}$ of the simulation. The output of the QSTS simulation may be a time series signal (e.g. phase voltages at a specific bus) or it may be a discrete value (e.g. total number of tap action for one year).

This dissertation focuses on five metrics although other metrics can be found using QSTS simulation. The voltage quality is determined using the highest and lowest feeder voltages and time spent outside voltage limits set by the American National Standards Institute (ANSI) [40]. The number of tap actions and capacitor switches are determined for three voltage regulating transformers and a capacitor bank on the feeder. Finally, the real

power line losses are also measured as a performance metrics of the QSTS simulation. The high-level diagram in Figure 5 summarizes the QSTS simulation algorithm.

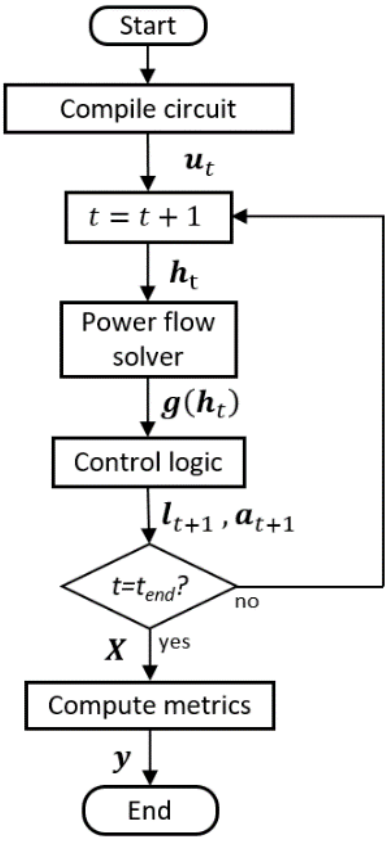


Figure 5 – Flow chart of the Brute force QSTS algorithm

While the input vector $\boldsymbol{u}_t$ is a predetermined time-series dataset, the vector $\boldsymbol{l}_t$ is time-dependent and can only be determined with the QSTS simulation. This is because voltage and current magnitudes from the unbalanced power flow solution are needed in the controller logic of the regulators and capacitors to determine whether an action should be taken before moving to the following time step. Since most distribution voltage controllers incorporate a delay, the general logic in Algorithm 1 below captures the temporal aspect of the simulation with a delay accumulator a_m for each controllable element m. $v_{control,m,t}$ is the control signal of element m at time t, $v_{ref,m}$ is the reference voltage setpoint of

control element *m*, $v_{band,m}$ is the controller bandwidth of element *m*, and δ_m is the delay of element *m*.

Algorithm 1: Voltage Control Element Switching Logic

$if\ |v_{control,m,t} - v_{ref,m}| \geq v_{band,m}/2:$
    $a_{m,t} = a_{m,t-1} + 1,$
    $if\ a_{m,t} < \delta_m:$
        $l_{m,t} = l_{m,t-1},$
    $elseif\ a_{m,t} \geq \delta_m:$
        $Controller\ action\ taken,$
        $Resolve\ power\ flow\ equations,$
    end
$else:$
    $l_{m,t} = l_{m,t-1},$
    $a_{m,t} = 0,$
end

At the end of each time-step, the controller logic determines if its control signal, based on the power flow solution, is inside the deadband. If it is, the controller state is the same as the one from the previous time-step and the delay accumulator is set to zero. On the other hand, if the signal is outside the deadband, the delay accumulator advances and if it exceeds the incorporated delay, a controller action is taken.

3.2 Proposed quantization algorithm

The brute force QSTS simulation repeatedly goes through solving the unbalanced three-phase power flow equations and checking whether an action is needed by any controllable element at each time step. While the controller logic is fast, non-linear power flow solvers can be computationally demanding especially for feeders with large number of buses. However, electrical demand is highly cyclical and similar power flows are bound to be repeated over a year-long simulation. The novel approach presented in this dissertation is to bypass the power flow computation if the solution for a given quantized

input has already been determined. The crucial contribution of this work is to define a power flow solution with respect to $\boldsymbol{h}_t$ and not $\boldsymbol{u}_t$ to handle regulators and capacitors with their controller logics. This is because there exist multiple valid solutions for a given load and PV power injection due to voltage-regulating equipment (see Challenge 4 in Section 2.2.4). At each time-step, $\boldsymbol{h}_t$ is compared to previous iterations to determine whether the power flow equations should be solved again. Equation (7) provides a high-level logic of this comparison.

$$\begin{gathered} If\ \boldsymbol{h}_t = \boldsymbol{h}_\tau \ \ s.t. \ \ \tau \in \{1, t\}, \\ then\ \boldsymbol{g}(\boldsymbol{h}_t) = \boldsymbol{g}(\boldsymbol{h}_\tau) \end{gathered} \tag{7}$$

This logic allows to reassign a power flow solution to a time-step without having to compute the power flow equations when a similar power flow has been computed. Such as the brute force approach, the proposed algorithm goes through each time step chronologically to capture the time dependence between time-steps. Figure 6 provides a high-level diagram of the proposed vector quantization algorithm.

Each unique $\boldsymbol{h}_t$ is a vector of discrete values that characterizes a unique power flow solution. The number of possible vector $\boldsymbol{h}_t$ can grow rapidly:

$$N_{\boldsymbol{h}_{possible}} = \prod_{i=1}^{N_{profile}} \eta_i * \prod_{i=1}^{N_{reg}} \rho_i \ * \prod_{i=1}^{N_{cap}} \kappa_i \ . \tag{8}$$

where η_i is the number of unique values in each profile i, ρ_i is the number of tap positions of regulator i, κ_i is the number of states of capacitor bank i, and $N_{profile}$, N_{reg}, and N_{cap} are the number of profiles (length of $\boldsymbol{u}_t$), regulators and capacitor banks in the circuit, respectively.

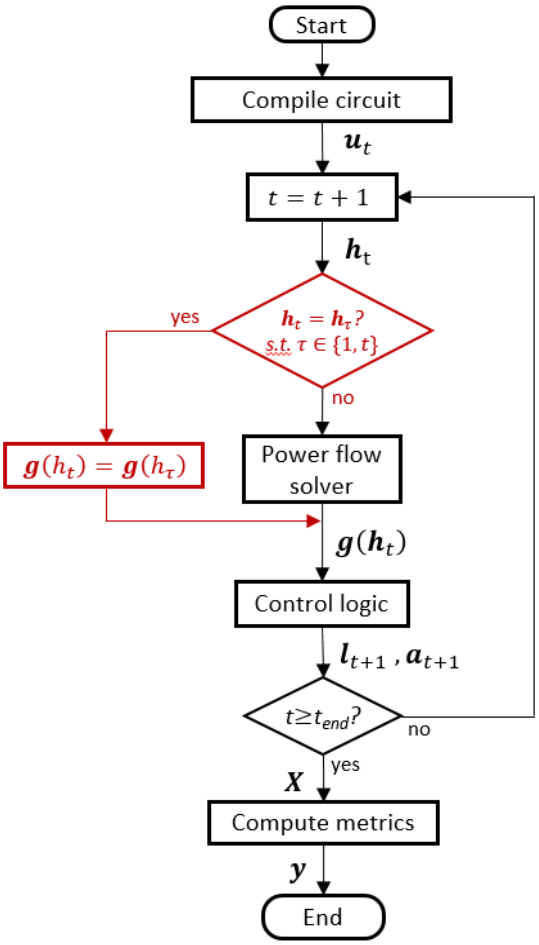


Figure 6 – Flow chart of the QSTS simulation with proposed quantization algorithm (in red).

Because the number of possible vector $\boldsymbol{h}_t$ can grow rapidly, it is important to implement the algorithm inline to only solve the states that are experienced over the time horizon. Additionally, the number of unique vector $\boldsymbol{h}_t$ can be reduced through vector quantization. Since the voltage regulator tap positions and capacitor statuses both have a relatively small set of discrete states and their impact on the voltage is significant, this information reduction technique is focused on quantizing the input vector $\boldsymbol{u}_t$, namely on the system demand and PV output profiles. As the number of unique vector $\boldsymbol{h}_t$ is decreased, the number of computed power flows is also reduced. An interesting research question to answer is how much error can be introduced by vector quantization before the accuracy of

the QSTS simulation ($\boldsymbol{y}$) becomes undesirable. A detail analysis of the correlation between the quantization error and the metric accuracy is discussed in details in Chapter 5.

3.3 Implementation of the algorithm

The vector quantization algorithm becomes relevant when the computational time of the quantization logic (red in Figure 6) is smaller compared to solving the power flow equations themselves. A pseudo-code of this quantization logic can be found in Algorithm 2. At the beginning of each time-step, the algorithm first determines whether a power flow with identical inputs ($\boldsymbol{h}_t$) has already been solved. This logic is implemented with matrix indexing (line 1) although any hashing functions mapping the inputs to a single variable can be used based on the scripting environment. This single variable (*indx*) represents the row index in the solution space $\boldsymbol{S}$ where the power flow solution $\boldsymbol{g}_t$ corresponding to $\boldsymbol{h}_t$ is stored (line 6-8). More specifically, the row index (*indx*) in $\boldsymbol{S}$ for that unique $\boldsymbol{h}_t$ is stored in a large indexing matrix $\boldsymbol{M}$ with the number of dimensions in $\boldsymbol{M}$ equal to the length of vector $\boldsymbol{h}_t$. Thus, an if-statement determines that a power flow should be computed when the hashing function returns a zero value and assigns the previously-computed power flow solution when a row index is returned. This allows $\boldsymbol{h}_t$ to be the multi-dimensional location of the index value in $\boldsymbol{S}$ without having to search $\boldsymbol{S}$ for $\boldsymbol{h}_t$. Because the solution space is extremely sparse, the algorithm is implemented inline and each new computed power flow solutions are appended to the solution space $\boldsymbol{S}$ (line 6). Furthermore, the $indx$ variable can be stored as a time series vector ($\boldsymbol{indx}$) to reconstruct any time series measurements on the feeder using the solution space matrix $\boldsymbol{S}$. This data compression capability is discussed in Section 3.6.2.

Algorithm 2: Quantization logic using matrix indexing

1: $indx = \boldsymbol{M}([\boldsymbol{h}_t])$,
2: $if\ indx \neq 0\ :$
3: $\quad \boldsymbol{g}_t = \boldsymbol{S}(indx, :)$
4: $else$:
5: $\quad Solve\ power\ flow\ equations\ with\ \boldsymbol{h}_t$,
6: $\quad indx = size(\boldsymbol{S}, 1) + 1$,
7: $\quad \boldsymbol{S}(indx, :) = \boldsymbol{g}_t$,
8: $\quad \boldsymbol{M}([\boldsymbol{h}_t]) = indx$,
9: $\quad \boldsymbol{indx}(t) = indx$,
10: end

Since the additional logic (line 1-2) is performed at all time-steps regardless of whether the power flow is solved, the computational time of assigning a solution (line 3) must be significantly smaller than solving the non-linear power flow equations (line 5) to account for this additional logic at each time-step. Because the algorithm is tested on a small circuit (see next section), this matrix indexing logic must be extremely fast compared to the iterative power flow solver to achieve attractive computational time reduction. The speed of the algorithm is discussed in Section 3.5.1.

3.4 Test case: modified IEEE 13 bus circuit

The vector quantization algorithm is first tested on a modified IEEE 13-bus test circuit that incorporates a centralized PV system at the end of the feeder [7], [41] (Figure 7). The circuit has three single-phase voltage-regulating tap changers, one three-phase and one single-phase capacitor bank. A voltage-based controller is introduced for the three-phase capacitor bank to study the impact of vector quantization on the number of switches performed by the capacitor bank. A single load profile is uniformly assigned to all loads as a scalar multiplier of their respective real and reactive peak power (Figure 8). Another profile is assigned to the PV system also as a per-unit multiplier of its rated power output (Figure 9).

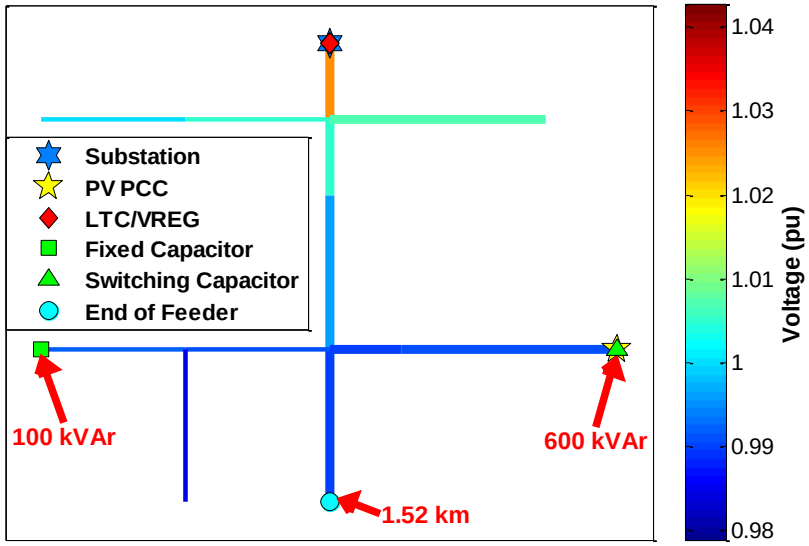


Figure 7 – Diagram of the modified IEEE 13-node feeder colored by voltage.

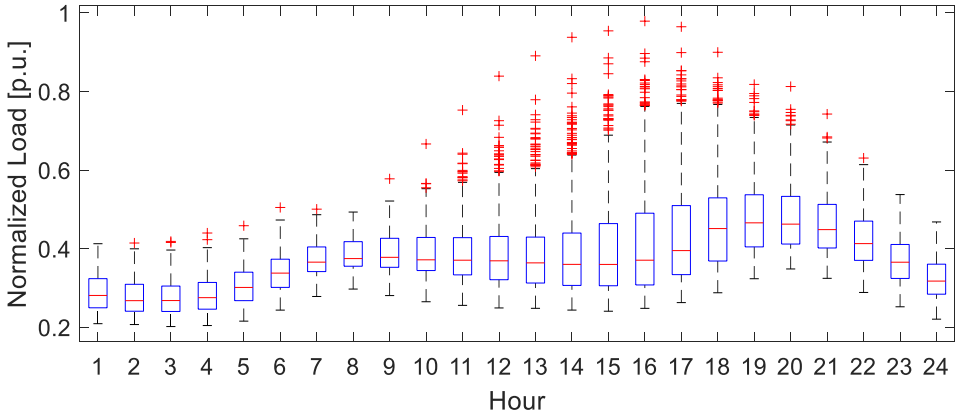


Figure 8 – Hourly box plot of the normalized load profile used in the modified IEEE 13 bus test case.

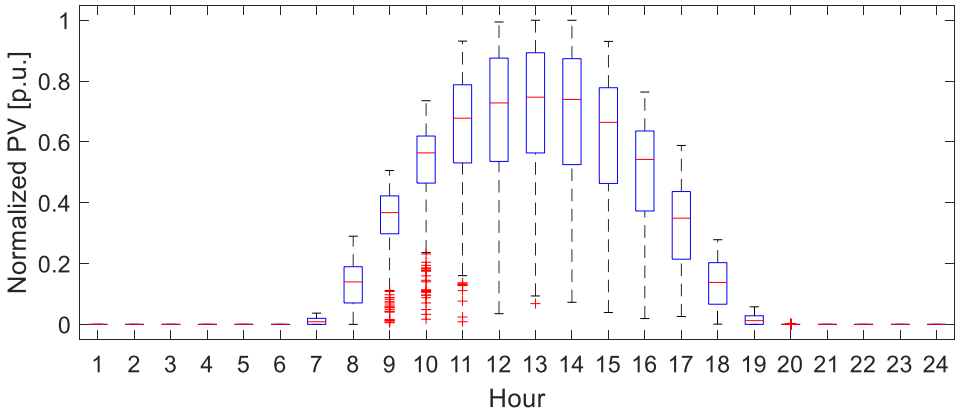


Figure 9 – Hourly box plot of the normalized PV profile used in the modified IEEE 13 bus test case.

The circuit is modeled in OpenDSS [33], and the algorithm is coded in Matlab using the GridPV toolbox [42] to interact with OpenDSS. Realistic dataset for the two profiles were obtained from normalized substation SCADA measurements in California and from global horizontal irradiance (GHI) data at the NREL Oahu site [43]. The yearlong dataset for the load profile is linearly extrapolated between the 5-minute measurements to generate the 1-second resolution profile. The GHI was sampled at 1 Hz and converted to PV power output for a latitude-tilt fixed PV system using the DIRINT decomposition model, Hay/Davies transposition model, and the Sandia Array performance model [44]. The precision of the magnitude in both normalized profiles is 10^{-5} p.u. As the precision of the profiles is reduced through vector quantization, the number of unique power flow solutions computed decreases. A heat map of the number of time steps with identical $\boldsymbol{u}_t$ profiles (Figure 10) provides a glimpse of the sparsity of the state space – 10.3% of $\boldsymbol{u}_t$ vectors are not experienced. Note that each color pixel may represent multiple unique power flow solutions since the regulator tap positions and capacitor status could not be represented (6 dimensions). Based on the figure, some load and PV combinations are experienced more than 10,000 times over the yearlong simulation justifying computing the power flow solution the first time the combination is experienced and reassigning that solution to any subsequent times it is experienced.

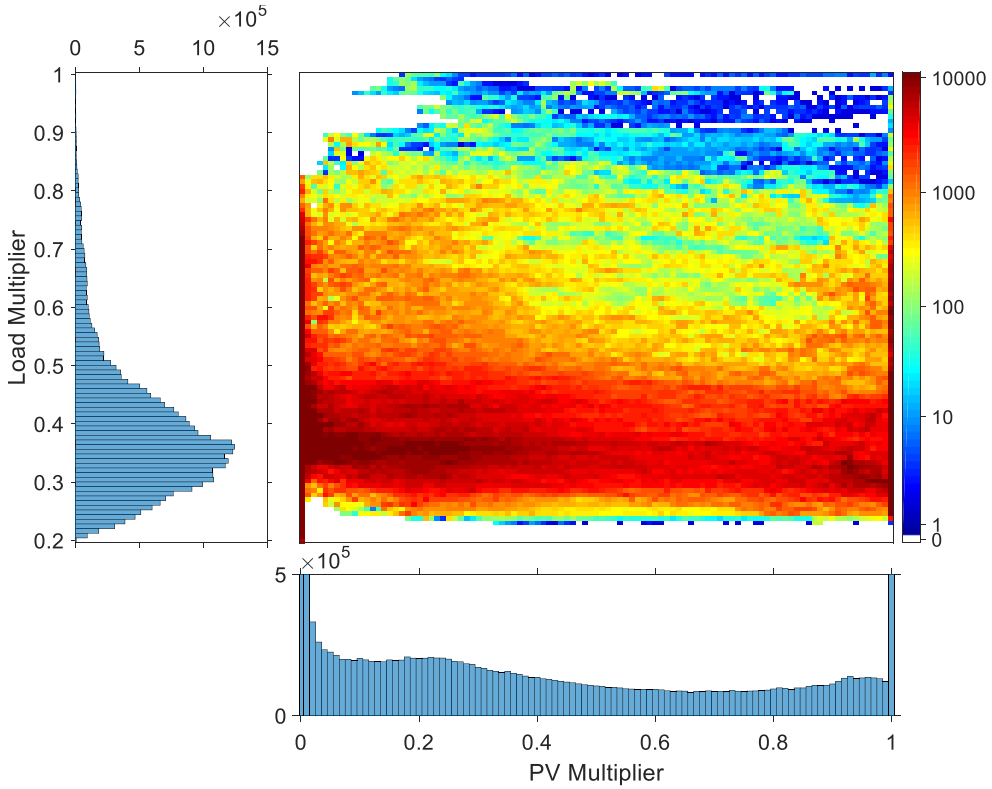


Figure 10 – Heat map of the number of time steps in a year with similar load and PV multiplier (u_t) when quantized in 101 clusters

3.5 Simulation results

3.5.1 Computational time reduction results

Table I shows the results of the QSTS simulation using the proposed quantization algorithm. The objective of a QSTS simulation is to determine how a DER interconnection would affect the operation of a distribution network feeder. In this work, the performance of the proposed vector quantization algorithm is tested based on how accurately, relative to the brute force time-series results, the algorithm can report the following metrics: number of controller actions (i.e. tap changes or capacitor switches), highest and lowest feeder voltage, total power loss, and time outside ANSI voltage limits [40]. In this work, the ANSI Range A standards (126/117V *on* $120V_{base}$) were applied to the medium voltage. Table I shows that the load and PV profiles are uniformly quantized into 51, 101 and 201

clusters over their range. The number of clusters were arbitrarily chosen based on the accuracy of the results but a vector quantization strategy is proposed in Chapter 5. The number of unique vectors $\boldsymbol{h}_t$ ($N_{h_{possible}}$) can be determined given the discrete set of values for the load and the PV profiles. Because $N_{h_{possible}}$ in each vector quantization case exceeds the number of time-steps, it is important to implement the algorithm inline and populate the solution space only with the power flow solutions that are computed ($N_{p.f.\ solved}$). By comparing the two, the sparsity of the solution space is less than 0.02% for each vector quantization case considered this study. Moreover, when compared to the number of computed power flows in the brute force, this sparse solution space is still less than 0.7% in each quantization case. For this specific reason, significant computational time reduction can be achieved by avoiding solving similar power flows. The algorithm is coded in MATLAB interacting through a COM object with OpenDSS to solve the power flow and the reduction in computational time with respect to running the brute force QSTS simulation is presented in the following table.

TABLE 2. Characteristics of the QSTS simulations with vector quantization

# of clusters (load,PV)	$N_{h_{possible}}$	$N_{p.f.\ solved}$	$\frac{N_{p.f.\ solved}}{N_{h_{possible}}}$	$\frac{N_{p.f.\ solved}}{31{,}536{,}000}$	Reduction in Comp. Time
Brute Force	31,536,000	31,536,000	100 %	100 %	-
Cluster 201,201	2,903,781,474	210,031	0.0072%	0.668%	98.90%
Cluster 201,101	1,459,114,074	119,305	0.0082%	0.379%	99.19%
Cluster 201, 51	736,780,374	66,557	0.0090%	0.212%	99.35%
Cluster 101,201	1,459,114,074	127,030	0.0087%	0.404%	99.16%
Cluster 101,101	733,186,674	70,466	0.0096%	0.224%	99.34%
Cluster 101, 51	370,222,974	38,783	0.0105%	0.123%	99.44%
Cluster 51,201	736,780,374	76,837	0.0104%	0.244%	99.32%
Cluster 51,101	370,222,974	42,256	0.0114%	0.134%	99.43%
Cluster 51, 51	186,944,274	23,109	0.0124%	0.074%	99.49%

Theoretically, the reduction in computational time of this algorithm should be equal to the reduction in the number of computed power flows. However, there are other computational times that affect the performance of the algorithm. Specifically, the total

computational time ($T_{comp.}$) is a function of the quantization logic (T_{QL}) in *Algorithm 2*, controller logic (T_{CL}) in *Algorithm 1*, and the power flow solver as a function of number of power flow computed $f(N_{pf\ solved})$.

$$T_{comp.} = T_{QL} + T_{CL} + f(N_{pf\ solved}), \tag{9}$$

where,

$$f(N_{pf\ solved}) = \left(\frac{T_{brute\ force}}{31.5e^6}\right) N_{pf\ solved}. \tag{10}$$

By testing the performance of the algorithm on a small test circuit, significant computational time reduction can be challenging. Since the power flow solver is extremely quick, T_{QL} and T_{CL} have a significant impact on the total computational time, hence the mismatch between the reduction in time and in the number of computed power flows (last two columns in TABLE 2). However, the computational time of a single power flow is proportional to the number of nodes in the system [30]. For example, the computational time of the brute-force QSTS simulation in OpenDSS for the IEEE 123-bus test circuit is 6.6 times longer, which is the same ratio as the number of nodes between both circuit (42 vs. 278 nodes). Because T_{QL} and T_{CL} are independent of the number of nodes in the feeder, the reduction in computational time will converge to the reduction in the number of computed power flows as the size of the feeder increases:

$$\frac{T_{comp.}}{T_{brute\ force}} = \frac{N_{pf\ solved}}{31.5e^6}. \tag{11}$$

Thus, the speed of the algorithm is directly dependent on the reduction in the number of power flow solved. Moreover, vector quantization purposely clusters similar power flows

to reduce the number of power flows and achieve attractive computational time. The strategy in a clustering method to reduce the individual power flow solutions is discussed in Chapter 5.

The discussion above excluded the computational time associated with the communication between MATLAB and OpenDSS. Since these COM calls would not be present if the algorithm were directly implemented in a commercial software, their computational time are substracted in the reported results by running the QSTS simulation with the solved $\boldsymbol{M}$ and $\boldsymbol{S}$ matrices. This helps determine the overhead computational time ($T_{QL} + T_{CL}$) of the algorithm by removing $f\left(N_{pf\ solved}\right)$ and its associated COM calls. Simulation results supporting this justification can be found in Appendix B. The computational time of the power flow solver ($f\left(N_{pf\ solved}\right)$) can then be added based on the number of power flow solved and the computational time of the brute force QSTS simulation as shown in Equation (10).

3.5.2 Quantization accuracy results

This fast time-series approximation has shown very attractive percent reduction in computational time considering the accuracy in all metrics (TABLE 3). Acceptable accuracy thresholds are established based on feedback from distribution system engineers as part of a broader project with Sandia National Laboratory. Each metric in TABLE 3 performs differently under vector quantization. Static metrics such as total line losses and highest/lowest feeder voltage can be determined with negligible error for each quantization level tested in this chapter. On the other hand, time-dependent metrics such as controller actions and time outside ANSI are more sensitive to the vector quantization. This is due to

the control algorithm being non-continuous (deadbands and delays) which amplifies the introduced quantization error. A small error in the controller voltage could trigger an action that would have not occurred otherwise. Furthermore, the complex interaction between controllers can create a completely different course of actions and create significant challenges in reporting these time-dependent metrics. For additional information on controller interactions, refer to [30].

TABLE 3. Quantization accuracy for a yearlong QSTS simulation at various quantization levels

# of Clusters Load / PV	# of Tap Changes Reg1 / Reg2 / Reg3	# of Cap Switches	Highest / Lowest Feeder voltage	Total Line Losses	Outside ANSI Above / Below
Brute Force (Accuracy Threshold)	7014 / 7202 / 8411 (10%/reg)	2476 (20%/cap)	1.0607/ 0.9673 p.u. (0.005p.u.)	146.0 kWh (5%)	22.13h / 11.47h (5%)
Cluster 201,201	1.0%/ 0.6%/ 1.2%	0.8 %	<0.0001/ -0.0002 p.u.	0 %	-0.1% / 3.0%
Cluster 201,101	0.5%/<-0.1%/ 0.5%	-0.2 %	<0.0001/ -0.0002 p.u.	0 %	<-0.1% / -0.2%
Cluster 201, 51	1.2%/ 1.0%/ 1.0%	1.1 %	<0.0001/ <0.0001 p.u.	0 %	<-0.1% / -1.8%
Cluster 101,201	-0.4%/-1.2%/ -0.3%	-1.5 %	<0.0001/ -0.0006 p.u.	<-0.1 %	-0.8% / 0.8%
Cluster 101,101	-0.2%/-1.1%/ -0.5%	-0.3 %	<0.0001/ -0.0007 p.u.	<-0.1 %	-0.2% / 1.1%
Cluster 101, 51	0.2%/-0.2%/ -0.1%	-0.6 %	<0.0001/ -0.0004 p.u.	0 %	-0.2% / -0.8%
Cluster 51,201	-2.3%/ -5.1%/ 0.4%	-5.9 %	<0.0001/ -0.0006 p.u.	<-0.1 %	-0.7% / -3.3%
Cluster 51,101	-1.4%/ -4.2%/ 1.0%	-4.9 %	<0.0001/ -0.0007 p.u.	0 %	-0.1% / -2.4%
Cluster 51, 51	-1.9%/ -4.8%/ 0.3%	-5.7 %	<0.0001/ -0.0007 p.u.	0 %	0.7% / -5.0%

Whether the profiles are quantized 51/101 or 101/51, the number of possible power flow solutions ($N_{h_{possible}}$) remains the constant. However, the number of computed power flow solutions is lower when the load profile has fewer clusters. This trend is also valid for clustering 201/101 vs, 101/201 and 201/51 vs. 51/201. The accuracy of the metrics analyzed in this set of simulations does not have a direct relationship with the number of computed power flows. As a matter of fact, the simulations with higher resolution in the load profile than in the PV profile yields more accurate results than vice versa. This is due to their respective impact on the feeder. Again, the optimal strategies to quantize the input vector according to their impact on the feeder is investigated in Chapter 5.

3.6 Discussion

3.6.1 Computation of unique power flow solutions

As unique power flows are computed, the presumptive observation would be that the number of new unique power flow solutions decreases as the simulation progresses throughout the year. This theory is explored by plotting the number of new power flow solutions computed for each day of the year. Figure 11 shows a histogram of the number of new computed power flow solutions for every day of the year if the simulation started on January 1st.

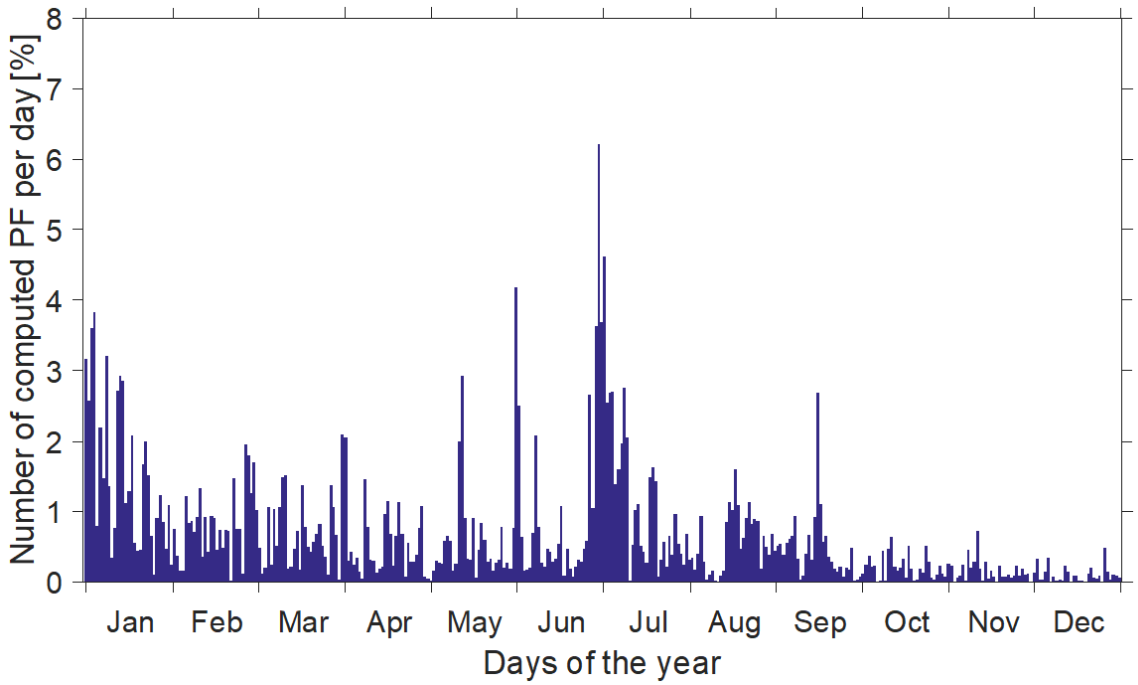


Figure 11 – Percentage of the new unique power flows computed daily.

In the first 4 months, a decaying trend is present where fewer new unique power flows are computed each day. However, a few spikes in new computed power flow solutions appear during the summer months. The load only approaches peak power during those summer months and causes new power flows to be computed. In Figure 12, a heat

map showing the number of new power flow solutions computed the first month of the simulation and during a summer month. This illustrates that a different region of the solution space is explored throughout the simulation. Again, note that multiple power flow solutions can have the same vector $\boldsymbol{u}$ because of controller states, hence some load/PV combinations (represented by pixels) can have 10 unique power flow solutions.

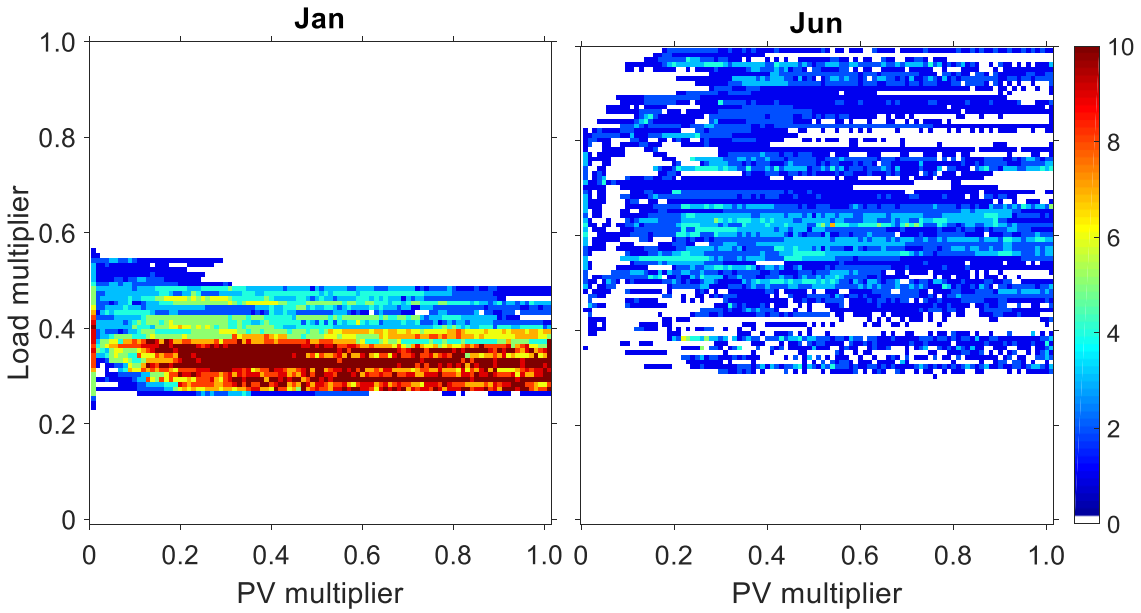


Figure 12 – Number of new computed power flows (g) with the same load and PV multipliers (u) between January and June.

Although starting the simulation during the summer month would result in computing more power flows upfront (Figure 13), the computational time reduction of the algorithm will remain the same since the same number of power flows are computed over the time horizon. Thus, the proposed algorithm is independent of which day it is started on as long as the controllers are started on the correct states.

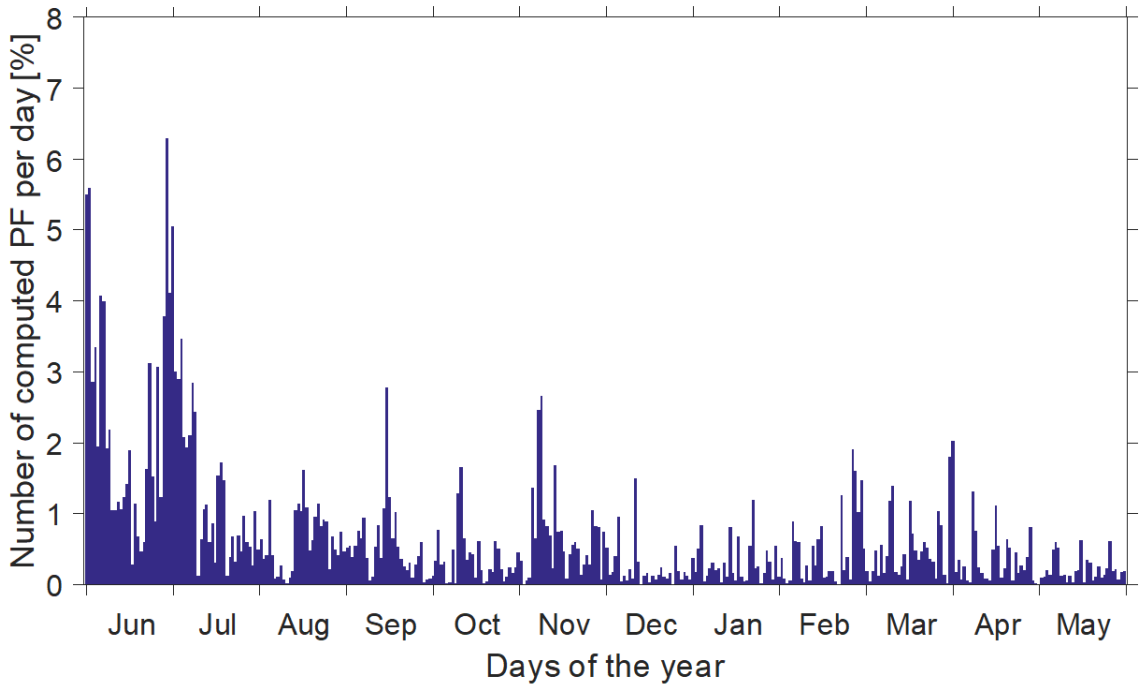


Figure 13 – Percentage of the new unique power flows computed daily when the simulation is started on June 1st.

3.6.2 Time-series data compression

One concern with conducting any time-series simulations is logging the various results from the analysis in matrix $\boldsymbol{X}$. For example, recording and storing multiple time-series datasets (such as voltage at each node on a feeder) for 31.5 million time-steps can easily exceed a computer's available memory. Thus, not all measurements can be recorded as time-series with current brute-force approaches. The proposed quantization algorithm provides real-time data compression for memory management by only storing unique solutions in the solution space matrix $\boldsymbol{S}$ and the solution index as a time-series vector ($\boldsymbol{indx}$). The time-series matrix $\boldsymbol{X}$ can easily be reconstructed ex-situ without having to run the QSTS simulation again to compute any metrics of interest.

$$\boldsymbol{X} = \boldsymbol{S}(\boldsymbol{indx}, :) \tag{12}$$

The data compression level is based on the number of unique power flow solved and is typically less than 1% of the total number of time steps. This is especially important if the interest of running the QSTS simulation is to capture the voltage variation at each buses of a large distribution system feeder. For instance, a feeder with 5000 nodes would require ~175GB of memory if solved with a brute-force method while the vector quantization algorithm would compress the data to less than 1.75GB. This represents a significant advantage for the proposed algorithm.

CHAPTER 4. VECTOR QUANTIZATION SCALABILITY

4.1 Introduction

The objective of developing a fast QSTS simulation algorithm is for the QSTS simulation to be widely used by the industry. Thus, it is crucial to develop a robust algorithm capable of simulating complex feeders. The vector quantization algorithm described in Section 3.2 is tested on a modified IEEE 13 bus test circuit, which raises the question of scalability to other IEEE or real distribution feeders. First, the number of buses is not representative of an actual feeder. Second, all loads and PV systems are each assumed to follow the same respective profiles. Third, only four controllable elements are considered, but some feeders have more devices. Last, controller logics specific to existing voltage-regulating equipment are modeled but more recent controllable devices, such as smart inverters, are not discussed. These concerns are addressed in this chapter to improve the robustness of the algorithm.

The vector quantization algorithm proposed in the previous chapter groups similar power flow solutions in clusters based on the factors that impact them to avoid the iterative power flow solver. Each time the power flow equations are solved, the solution is stored in a solution space for subsequent time steps. The power flow solution depends on two factors: the power injections on the feeder and the previous states of controllable elements. For PV interconnection studies, the power injections at a time *t* are defined as the vector $\boldsymbol{u}_t$ where $\boldsymbol{d}_t$ is a vector of the different load profiles and $\boldsymbol{p}_{pv,t}$ is a vector of the PV output profiles. The previous states of controllable elements at time *t* defined as $\boldsymbol{l}_t$ includes the

states of any elements that affect the voltage on the feeder, for instance any voltage-regulating tap changers $\boldsymbol{r}_t$ or capacitor banks $\boldsymbol{c}_t$. The factors affecting the power flow solution can then be defined as vector $\boldsymbol{h}_t$. The quantization logic used to determine whether a solution exists must be faster than solving the power flow equations to make the algorithm attractive. The logic discussed in the previous chapter uses a matrix indexing method as shown in Algorithm 2. Matrix $\boldsymbol{S}$ is the solution space in which solutions are appended each time a new solution is computed. Matrix $\boldsymbol{M}$ is an indexing matrix used to determine whether or not a solution exists. Each value in vector $\boldsymbol{h}_t$ is associated to a unique dimension in the matrix $\boldsymbol{M}$. Thus, for an *n*-dimensional matrix $\boldsymbol{M}$, $indx = \boldsymbol{M}([\boldsymbol{h}_t])$ is the unique value at the location $\boldsymbol{M}(h_{t,1}, h_{t,2}, \dots, h_{t,n})$ that is associated with a power flow solution.

This matrix indexing logic runs quickly in MATLAB if the two matrices are pre-allocated in the memory. However, the indexing matrix $\boldsymbol{M}$ can become extremely large considering that it has the same number of dimensions as the length of vector $\boldsymbol{h}_t$, which is the total number of input time-series profiles and the number of controllable elements. The length of each dimension of $\boldsymbol{M}$ depends on the number of unique combinations of vector $\boldsymbol{h}_t$, which can be decreased using vector quantization. However, quantization inherently introduces an error in the data that can affect the accuracy of the simulation. Namely, metrics reported by the QSTS simulation, such as controller actions, under-/over-voltage, power loss, or constraint violations, can be falsely reported when quantization clusters a wide range of slightly similar power flow solutions. Because the algorithm requires the indexing matrix $\boldsymbol{M}$ to be pre-allocated in the memory, a finite number of load and PV profiles and controllable elements can be modeled based on the available memory for this multi-dimensional matrix. The formulation of this algorithm in the previous chapter limited

the combined number of profile and controllable elements to 6. These issues are addressed in the following section.

4.2 Scalability concerns of the proposed VQ algorithm

4.2.1 Size of the feeder

The test circuit simulated in the previous chapter has only 13 buses and thus, is not representative of a realistic distribution feeder with thousands of buses. Although circuit reduction algorithms have been proposed in the literature to accurately model a large distribution system feeder based on a small number of buses [39], the vector quantization algorithm scales well with the size of the feeder. As the number of nodes in the circuit increases, the computational time in OpenDSS required to solve an individual power flow increases proportionally [30]. The computational time of the proposed algorithm is a function of the number of power flows computed ($N_{pf\ solved}$), the controller logic (T_{CL}), and an overhead time associated with the implementation of the algorithm itself (T_{VQ}). The average computational time to solve a power flow can be estimated from the computational time of QSTS simulation with a brute force approach ($T_{brute\ force}$) by combining Eq. (9) & (10).

$$T_{comp.} = \left(\frac{T_{brute\ force}}{31.5e^{6}}\right) N_{pf\ solved} + T_{CL} + T_{VQ} \tag{13}$$

Because the quantization algorithm only computes power flows for a subset of the total number of time steps, the overall computational time of the vector quantization will increase much slower than if the power flow at each time step was computed. Furthermore, the overhead computational time associated with the controller logic and the logic required

to implement the algorithm is independent of the number of nodes in the circuit – the size of $\boldsymbol{h}_t$ and $\boldsymbol{M}$ in line 1 in Algorithm 2 does not change based on the number of nodes in the circuit. Therefore, as the feeder size increases, this overhead computational time becomes negligible compared to the computational time of the power flow solver, making the computational time reduction equal to the ratio of the number of power flows computed over the total number of time steps. For instance, a simulation with 31,500 unique computed power flows would reduce the computational time of a yearlong QSTS simulation by about a thousand.

4.2.2 Number of Load/PV Profiles

In the modified IEEE 13-bus test circuit discussed in Chapter 3, the indexing matrix $\boldsymbol{M}$ has 6 dimensions with two input time-series profiles, three voltage regulators, and one switching capacitor bank. If the two profiles are quantized into 100 clusters, the regulators have 33 tap positions, and the capacitor bank has two states, the indexing matrix $\boldsymbol{M}$ would have 719 million entries. Obviously, additional input time-series profiles for each type of customer (residential, commercial, industrial, etc.) or different PV locations cannot be modeled with the current algorithm. Each additional input time-series increases the size of $\boldsymbol{M}$ exponentially. In order to address this issue, the proposed method does not treat profiles as individual dimensions in the indexing matrix $\boldsymbol{M}$, but as a single dimension representing a 'scenario' of profiles. A time-series vector $\boldsymbol{indx_u}$ can be created where each entry is the index of the first time that combination of profiles was experienced. Furthermore, any combinations not experienced over the time horizon would not have a placeholder in the indexing matrix $\boldsymbol{M}$. In other words, the proposed method takes advantages of the sparsity

of the space to reduce the memory requirement to pre-allocate the indexing matrix. This sparsity is demonstrated in Section 4.3.

After quantizing the profiles, the time-series vector $\boldsymbol{indx_u} \in \mathbb{R}^{T\times 1}$ can be created representing the 'scenario' of profiles at each time-step. The value of each profile at a specific time step *t* can easily be determined with $\boldsymbol{u}([\boldsymbol{indx_u}])$. The vector defining unique power flow solutions becomes:

$$\boldsymbol{h}_t = \left[indx_{\boldsymbol{u},t}, \boldsymbol{l}_t\right] \tag{14}$$

By pre-processing the profiles, the number of dimensions for profiles in the indexing matrix is limited to one no matter how many input time-series profiles are considered. Moreover, the size of the indexing matrix is also reduced since not every unique vector $\boldsymbol{u}_t$ is experienced over the time horizon. This method is possible only because the dataset for the profiles is known prior to the time-series simulation. With this indexing method, the vector quantization algorithm is not affected by the number of profiles considered in the simulation, whether it is load profiles based on customer types or multiple PV profiles based on their scale and location. Note that quantization of the profiles is still required to achieve attractive computational time reduction. This is discussed in Chapter 5.

4.2.3 Number of Controllable elements

A similar approach to the one discussed above can be used to reduce the memory requirement for the controllable elements in the indexing matrix. However, unlike the profiles, the states of those elements are not known prior to the time-series simulation. Thus, the entire space must be created for the current indexing method to work. From

analyzing the simulation results on the modified IEEE 13-bus test circuit from Section 3.4, one can find that this space is extremely sparse as shown in Fig. 2.

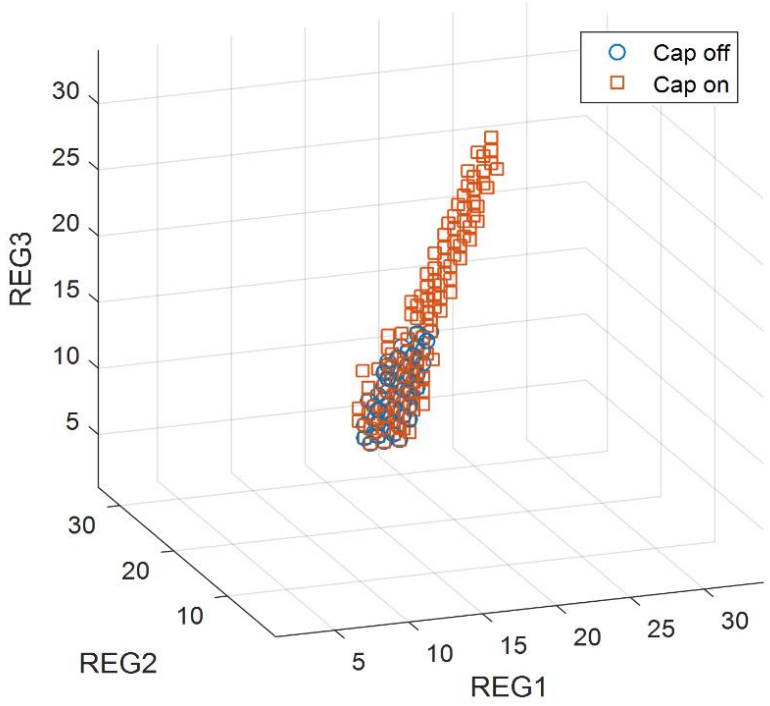


Figure 14 – Regulator tap position and capacitor status combinations that are experienced in a yearlong QSTS simulation for the modified IEEE 13-bus test circuit.

In order to take advantage of this sparsity, a controllable element matrix ***L*** can be created with each dimension representing the state of a controller. This matrix is initially a zero-matrix that is populated with indices as the simulation progresses. Each time a combination of controllable element states is experienced for the first time, a column is appended to the indexing matrix ***M*** and an index referring to the column number is stored in matrix ***L***. Because the number of combinations of controllable element states is not extremely high (250 times in the modified IEEE 13 bus test case), appending a column to the indexing matrix does not impact the computational time of the algorithm. The following two equalities are then implemented before line 1 of Algorithm 2.

$$indx_{\boldsymbol{L}} = \boldsymbol{L}([\boldsymbol{l}_t]) \tag{15}$$

$$\boldsymbol{h}_t = [indx_{\boldsymbol{u}}, indx_{\boldsymbol{L}}] \tag{16}$$

At the end of the yearlong simulation, matrix $\boldsymbol{M}$ has a number of rows equal to the number of profile scenarios and number of columns equal to the number of controllable element state combinations experienced. This formulation takes advantage of the sparsity of the space due to the correlation between controllable elements and possibly reduces the pre-allocated space for the controllable elements in the indexing matrix. In the 13-bus test case referenced above, the three phases are relatively balanced, explaining why the voltage regulator tap positions of each phase are highly correlated. Not every entry in matrix $\boldsymbol{L}$ needs to have a value referring to the indexing matrix $\boldsymbol{M}$.

This matrix decomposition method reduces the memory requirement to pre-allocate large datasets. This concept can be further expended as the number of controllable elements increases, which increases the size of $\boldsymbol{L}$ as well as its sparsity. Furthermore, the computational time of the algorithm will not be affected since Eq. (15) is only performed when a controller action is performed.

4.2.4 Types of controllers

Distribution feeders can have a wide variety of devices with controller logic: energy storage systems, smart inverters, regulators, capacitor banks, etc. The number of devices implemented on a feeder can rapidly grow as some of them are often installed with other DERs (e.g. advanced inverters with PV systems). In fact, the need to determine the settings for these new types of controls is an important application of QSTS simulation. For example, in order to determine adequate settings for energy storage discrete controllers

[46] or advanced inverter controllers [47], potentially thousands of QSTS simulations need to be performed with different settings to fully study any potential interactions between controllers and potential benefits.

One of the advantages of the proposed quantization algorithm is that it is robust to any type of controller logic. Generally speaking, discrete controllers operate based on a specific signal, whether it is time, voltage/current magnitudes, power factor, price, etc. The type of signal will dictate how the controller logic is modeled for a QSTS simulation and thus implemented in the vector quantization algorithm. For this purpose, discrete controllers can be grouped into three categories based on the input signal they require.

First, most controllers are dependent on a signal derived from the power flow solution (i.e. voltage/current magnitude, power factor, …) and have a hysteresis that prohibits them to be computed ahead of the time-series simulation. For instance, capacitor banks can operate based on a voltage signal and have delays and deadbands to reduce excessive operation. To accurately model these types of controllers in a QSTS simulation, their logic is implemented at the end of each time step. States from the power flow solution can be used as control signals and controller states (i.e. equipment states or delay accumulators) can model the hysteresis within the controllers. In the vector quantization algorithm, these equipment states (e.g. tap position) are referred to in the matrix $\boldsymbol{L}$ since they affect the power flow solutions. Because the algorithm goes through each time-step, delay accumulators and previous controller states can easily keep track of any expiring delays within the controllers as the simulation progress through time. As discussed in Section 4.2.3, the algorithm is not limited by the number of these controllers in the feeder model.

Second, the controllers that operates based on the power flow solution but does not have a hysteresis (e.g. deadbands, delays, state-of-charge, etc.) will always have the same outcome for a given power flow solution and system state. Thus, their controller logic can be implemented separately into the power flow module, with the vector quantization algorithm only storing the final solutions. Because each power flow is defined by a unique vector $\boldsymbol{u}$, the algorithm does not need to go through that controller logic each time a power flow has already been computed since only one output from the controller is possible for a given vector $\boldsymbol{u}$. For example, this subtle difference is especially advantageous for feeders with multiple advanced inverters with VOLT/VAR capabilities programmed to control the power factor of their output based on the voltage magnitude of the system. In a QSTS simulation, smart inverters are modeled with a controller logic iteratively changing the output of the inverter and re-computing the power flow until it converges since the outcome of the controller impact the local voltage that is used as an input signal in the controller. Under a brute-force approach to the QSTS simulation, this would be done at each time step, which drastically impacts the computational time. However, the proposed algorithm reduces the number of times it goes through this type of controller logic since only the final converged solution is stored by the vector quantization algorithm the first time that power flow solution is computed.

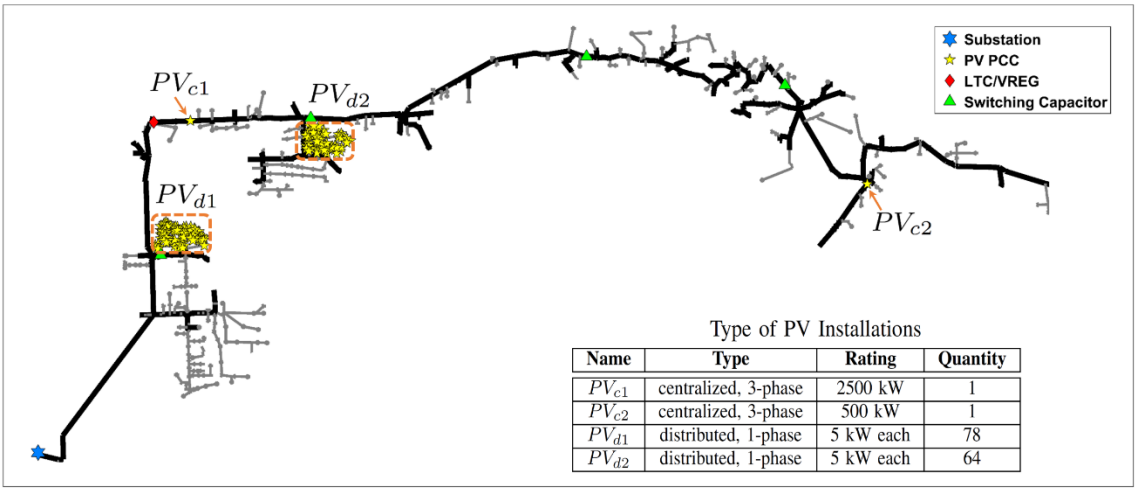
Third, any controllers operating solely on a signal known ahead of the QSTS simulation (i.e. price, time, etc.) and independent on any states from the power flow solution could be preprocessed with the other input time-series profiles as a scenario (see Section 4.2.2). For instance, the power output of an energy storage system or the charging of an electric vehicle can be modeled based on a predetermined schedule. In that case, the

power output profile is time-dependent and is treated similar to a PV or load profile in the input vector $\boldsymbol{u}$.

Thus, the proposed vector quantization algorithm is robust to different types of controllable elements that would be modeled with QSTS simulations since most discrete controllers would fit in one of the three aforementioned categories.

4.3 Test case: large realistic distribution system feeder

The scalability of the algorithm is demonstrated using an actual distribution feeder with 2969 buses (5469 nodes) and 9 controllable elements – 3 three-phase and 2 single-phase switching capacitor banks, 3 single-phase line voltage regulators, and a three-phase substation load tap changer (Figure 15). In addition, 144 PV systems are introduced in the feeder model to emulate a test case where DER introduce a significant impact on the operation of the feeder.



Type of PV Installations

Name	Type	Rating	Quantity
PV_{c1}	centralized, 3-phase	2500 kW	1
PV_{c2}	centralized, 3-phase	500 kW	1
PV_{d1}	distributed, 1-phase	5 kW each	78
PV_{d2}	distributed, 1-phase	5 kW each	64

Figure 15 – Topology of the actual distribution feeder.

The loads are categorized into two types of customers: residential (single-phase) and commercial customers (three-phase). The 1131 residential customers are mostly on the

laterals (grey traces in Figure 15) and account for 4.2 MW of peak power. The 317 commercial customers account for 1.7 MW on the backbone of the feeder (black traces in Figure 15). Each load type is modeled with a different profile to capture the difference in load behavior throughout the day (Figure 16 & Figure 17).

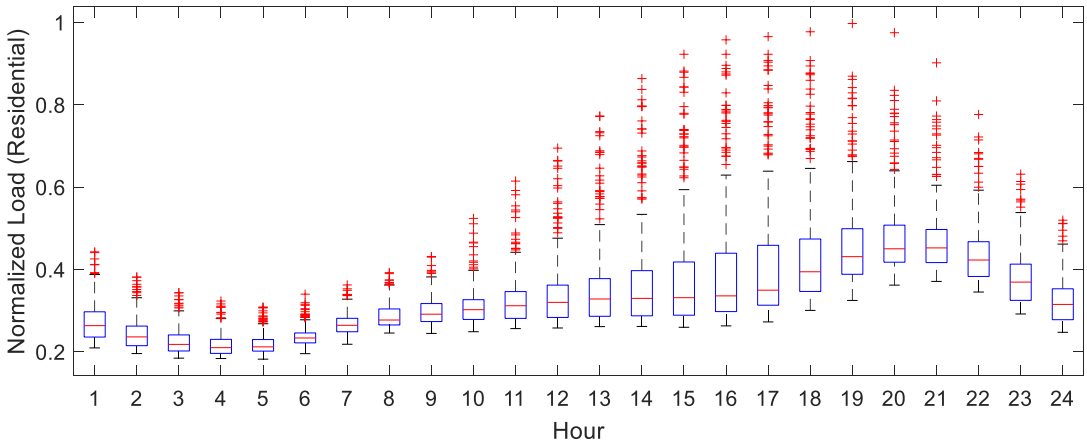


Figure 16 – Hourly box plot for the residential load profile demonstrating a peak during the late afternoon and early evening hours.

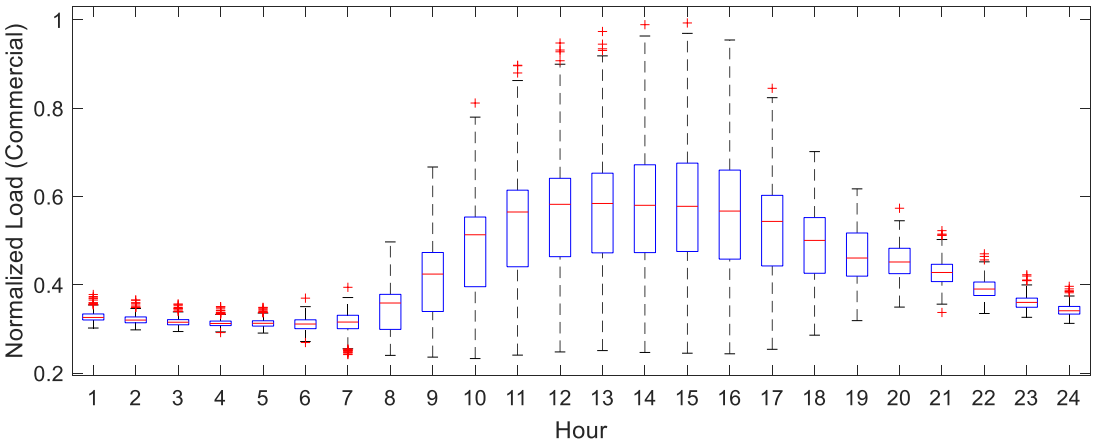


Figure 17 – Hourly box plot for the commercial load profile demonstrating a peak during business hours.

As expected, the residential profile tends to peak around 7p-8p while the commercial profile has a higher load during business hours. When both profiles are used to model the load on a distribution feeder, not every combination of load multiplier will be experienced (e.g. commercial load at 1 p.u. and residential load at 0.2 p.u.). Thus, this space

is relatively sparse and can be illustrated with a heat map of the reoccurrence of each combination (Figure 18). This further justify the discussion in Section 4.2.2.

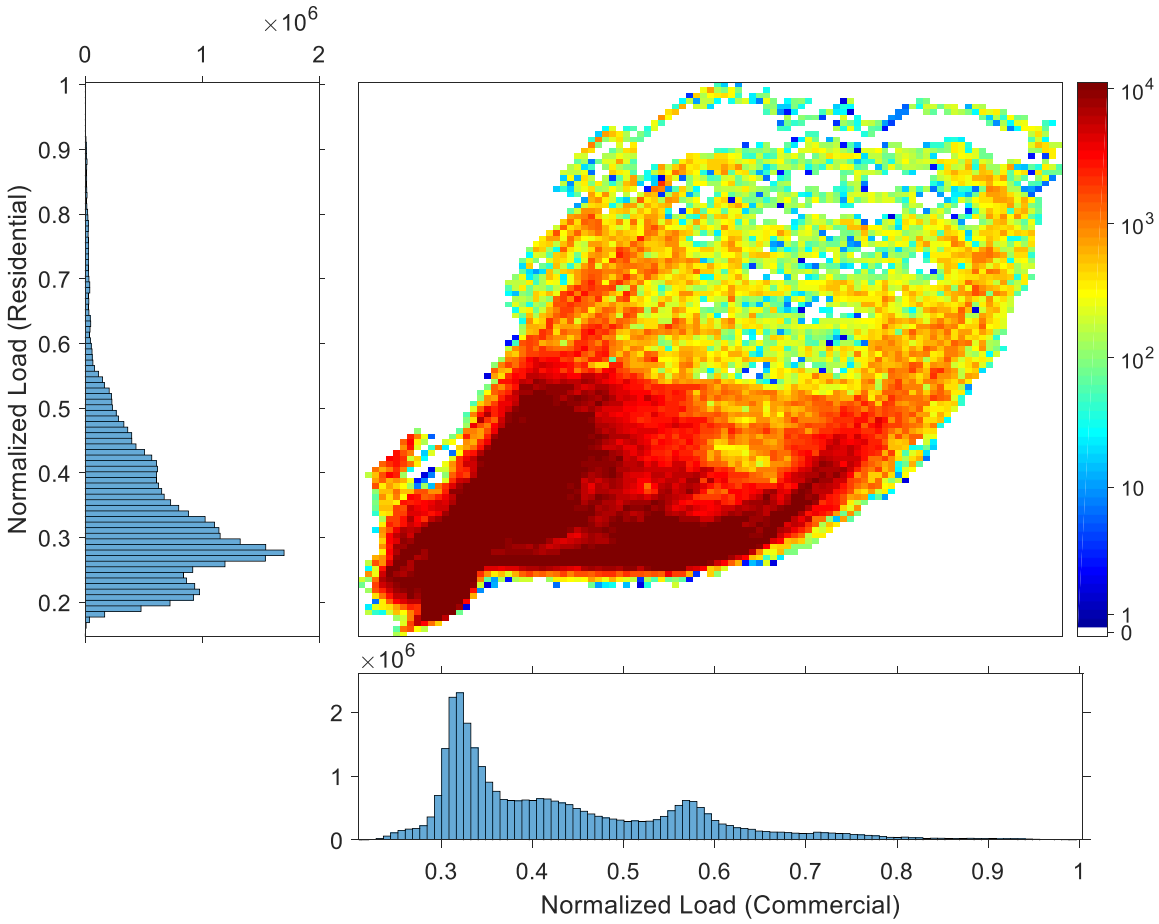


Figure 18 – Heat map of the number of time steps that have the same scenario (combination of multiplier values) for a yearlong profile with 1-second resolution. Note that the profiles were clustered for illustrative purposes.

The PV systems are classified into four regions to appropriately model their power output with individual profiles. Two centralized systems (PV_{c1} and PV_{c2} in Figure 15) are 2.5MW and 0.5MW respectively. All other PV systems are 5kW in size and are aggregated into two neighborhoods to model the impact of distributed PV systems and all systems in the same neighborhood are assumed to follow the same profile. Four synthetic PV profiles are created using a wavelet-based variability model presented in [48] at a wind speed of 20 m/s. Their box plots can be found in Appendix C.

4.4 Simulation results

The realistic distribution feeder test case discussed in the section above was created to test the scalability of the vector quantization algorithm. The test case has a large number of buses representative of a real distribution feeder, a total of 6 profiles (2 load and 4 PV profiles), and 9 controllable elements. Because of the increased number of profiles and controllable elements, the number of possible power flow solution has exponentially increased. For example, if each profile has 101 discrete values (clusters), the total number of possible power flows solutions is in the order of 10^{21} further justifying implementing the algorithm inline. Three simulations where each profile is quantized equally are carried out to demonstrate that the scalability of the algorithm is addressed (TABLE 4).

TABLE 4
Characteristics of the QSTS simulations with vector quantization

# of clusters (All profiles)	$N_{h_{possible}}$	$N_{p.f.\ solved}$	$\frac{N_{p.f.\ solved}}{N_{h_{possible}}}$	$\frac{N_{p.f.\ solved}}{31{,}536{,}000}$	Reduction in Comp. Time
Brute Force	31,536,000	31,536,000	100 %	100 %	-
Cluster (201)	2.50×10^{21}	4,127,077	~ 0 %	13.12 %	86.88 %
Cluster (101)	4.03×10^{19}	3,263,417	~ 0 %	10.35 %	89.65 %
Cluster (51)	6.68×10^{17}	2,030,666	~ 0 %	6.44 %	93.56 %

Because of the increased number of profiles compared to the modified IEEE 13 bus test case, the number of power flows solved at each quantization level has increased, which impacts the computational time. Note that the reduction in computational time is equivalent to the reduction in computed power flows since the overhead computational time (vector quantization logic and controller logic) are negligible (< 10 sec.) compared to the computational time of a power flow. The analysis on the factors impacting the computational time is discussed in detail in Section 3.5.

Because of the vector quantization, the accuracy of the reported metrics is impacted with respect to the brute force results. In TABLE 5, the error in the number of controller actions are reported by the proposed VQ algorithm at three quantization levels.

TABLE 5.
Quantization accuracy for a yearlong QSTS simulation at various quantization levels

# of Clusters	# of Tap Changes				# of Cap Switches				
All Profiles	**SUB**	**REG1**	**REG2**	**REG3**	**CAP1**	**CAP2**	**CAP3**	**CAP4-A**	**CAP4-B**
Brute Force	2746	5515	5055	5274	352	54	30	544	740
(Accuracy Threshold)	(10%)	(10%)	(10%)	(10%)	(20%)	(20%)	(20%)	(20%)	(20%)
Cluster (201)	0.58%	0.29%	0.46%	0.34%	2.27%	-11.11%	6.67%	-1.10%	0.27%
Cluster (101)	1.60%	2.65%	4.08%	3.60%	-1.14%	-11.11%	6.67%	-2.21%	1.89%
Cluster (51)	3.20%	6.13%	9.28%	5.50%	3.98%	-3.70%	-6.67%	1.47%	0.81%

Note that the percent error for CAP2 and CAP3 is higher because of the small number of actions recorded. From the perspective of distribution planning engineers, a difference of 30 or 32 controller actions over a yearlong period is not crucial and would result in the same conclusion that these controllers are not impacted by the solar PV systems. The accuracy of other controllers remained well within the accuracy thresholds. Thus, these results demonstrate that the vector quantization algorithm is scalable to realistic distribution feeders.

4.5 Smart inverter simulation

The last scalability concern with the vector quantization algorithm is feeders with different types controllable devices. In this section, the solar PV system in the modified IEEE 13 bus test case (section 3.4) is modeled with a smart inverter with VOLT/VAR capability. The inverter is controlled to change its power factor as a function of the voltage it sees at the point of interconnection. The following VOLT/VAR curve is specified in the model (Figure 19).

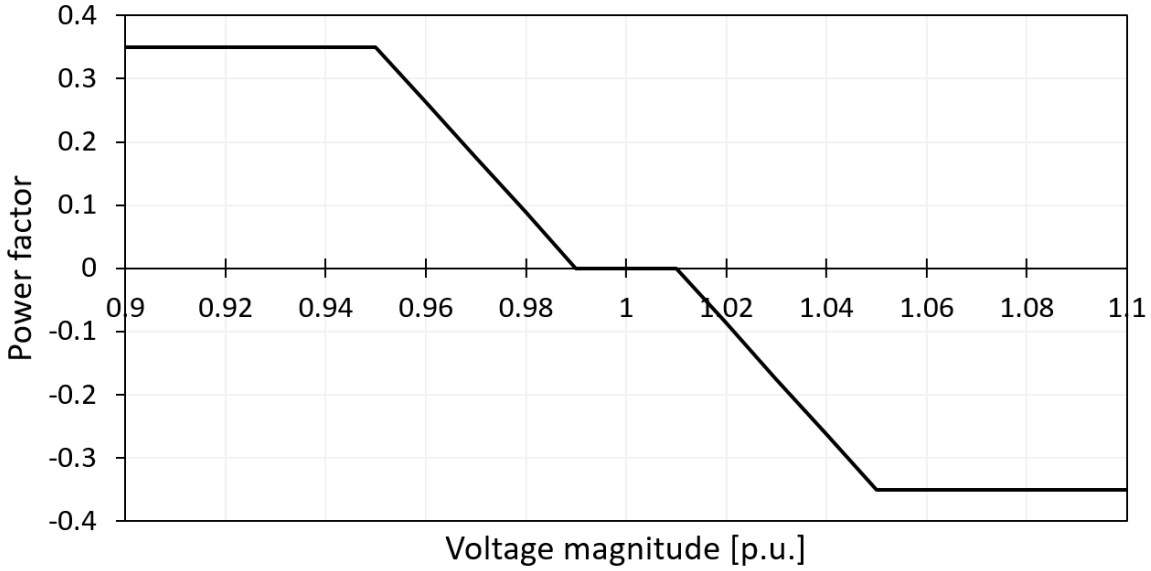


Figure 19 – VOLT/VAR curve of the smart inverter in the modified IEEE 13 bus test case.

Because the QSTS simulation solves static power flows, the controller for the smart inverter is slightly more complex to model and requires an iterative process in OpenDSS to determine the output power factor. The reason an iterative process is required is because a static power flow is solved assuming a given power factor by the inverter and based on the power flow solution (i.e. voltage magnitude and angle), the power factor may need to be updated which as a result will affect the voltage at the bus. Thus, the power flow solver and controller may iterate until the voltage and power factor converges. This iterative process has been studied in the literature [49] but is outside the scope of this work.

The vector quantization algorithm does not require any modifications to be capable of simulating feeders with smart inverters. Specifically, the algorithm only stores the power flow solution with the converged power factor. This is possible because the smart inverter controller should converge to the same power factor if the power injections are similar and the state of voltage-regulating equipment (capacitor banks, LTCs,…) are also identical. In other words, the controller of the smart inverter should converge to the same power factor

for a given vector $\boldsymbol{h}_{(t)}$. Thus, this type of controller does depend on the power flow solution but is not modeled with any deadband or delays that would require it to be implemented in the algorithm similarly to capacitor banks or tap changers.

Because of the additional power flows required to be solved at each time-step, the computational time of the QSTS simulation with the brute-force approach increases. Similarly, the vector quantization algorithm would also see an increase in the computational time to determine the power flow solution although this would only be the case when the power flow solution is computed and not at every time step. Thus, a significant reduction in computational time with the vector quantization algorithm is still possible. This is tested on the modified IEEE 13 bus test circuit with the smart inverter following the VOLT/VAR curve described above. The accuracy results are shown below for three quantization levels.

TABLE 6. Error in the reported metrics for the modified IEEE 13 bus circuit with a smart inverter (VOLT/VAR control)

# of Clusters Load / PV	# of Tap Changes Reg1 / Reg2 / Reg3	# of Cap Switches	Highest / Lowest Feeder voltage	Total Line Losses	Outside ANSI Above / Below
Brute Force (Accuracy Threshold)	2140 / 2168 / 3098 (10%/reg)	542 (20%/cap)	1.0575 / 0.9678 p.u. (0.005p.u.)	144.6 kWh (5%)	20.49h / 13.23h (5%)
Cluster 201,201	0.1% / -0.74% / -1.0%	-3.7%	1.0575 / 0.9674 p.u.	0.00 %	0.43% / -0.23%
Cluster 101,101	-3.8% / -7.7% / -1.8%	-24.4%	1.0575 / 0.9676 p.u.	-0.02 %	1.03% / 1.44%
Cluster 51, 51	-5.7% / -8.9% / 2.4%	-29.9%	1.0575 / 0.9671 p.u.	0.00 %	0.70% / -2.69%

These results demonstrate that the vector quantization algorithm can approximate the operation of various controllers on distribution feeders with smart inverters. Furthermore, the computational time reduction is not impacted since it is equal to the reduction in the number of computed power flow solutions and the error in the various reported metrics remains within the accuracy threshold. Note that the error observed in the number of actions in the capacitor bank is higher than the threshold for 101- and 51-clusters

simulations. This is due to the location of the capacitor bank being at the same bus as the smart inverter and due to the complex interaction between controllable elements. A single erratic tap action due to quantization will trigger a different course of action by other controllers and since the smart inverter also introduces a reactive power injection, the operation of the capacitor bank is impacted. Although a test case with a capacitor bank and smart inverter at the same point of interconnection is not realistic, it is an extreme case to test the efficacy of the algorithm. A strategy to properly quantize the profiles to minimize the error is discussed in Chapter 5.

4.6 Discussion

4.6.1 Number of load/PV profiles

Electric utilities commonly assign the feeder loads with load allocation [31] and then, scale all the loads based on a multiplier to achieve a desired total feeder load. In this chapter, the test case with the realistic distribution feeder is modelled with two load profiles based on number of phases of the load to simulate the behaviour of different customer types which provides a more detailed model of the power flow. One could argue that this is still an approximation of the real operation of a feeder since not every load is modelled with its own profile. Although this would provide a more accurate and detailed analysis of the operation of a distribution feeder, this cannot be done without high resolution data specific to a given distribution feeder, which is not currently available. The proposed algorithm addresses the current industry needs with the capability to simulate the operation of a distribution feeder based on available substation data (one profile) and the possibility to model the different types of customers if the data is available. Even though the current

algorithm scales to multiple profiles, a large number of profiles may create a large number of unique ‘scenarios’ which would impact the computational time reduction. As data will become available with the introduction of Advanced Metering Infrastructure (AMI) data, additional profiles could be introduced into the feeder models but this is subject of future work due to the limitation of the available data.

4.6.2 Significance of a scalable algorithm

In this chapter, the scalability of the vector quantization algorithm proposed in Chapter 3 is addressed in terms of the size of the feeder, the number of power injection profiles, the number of controllable devices, and the type of controllers that could be modeled in a QSTS simulation. The objective of this chapter is to insure the proposed algorithm is agnostic to the type or complexity of a distribution network feeder that would be modeled with QSTS simulations. The algorithm is verified on another test case with an increased number of buses (5000+), controllable elements (9), and load/PV profiles (6). Furthermore, the algorithm is tested on a feeder with smart inverters with VOLT/VAR capabilities. In both cases, results show that the algorithm is scalable and capable to simulate feeders with multiple degrees of complexities.

The significance of the work presented in this chapter is that the algorithm is its practicality and the possibility of it to be implemented into a commercial-grade distribution planning software. As project partners, OpenDSS and CYME have expressed interests in implementing the proposed VQ algorithm into their software because of the various benefits it provides. Namely, the versatility of the algorithm (Chapter 4) and its capability of time-series data compression (see Section 3.6.2).

In this chapter, the number of power injection profiles was addressed by reducing the number of degrees of freedom to one and treating the power injections as a ‘scenario’ of profiles. This drastically reduces the size of the indexing matrix but it does not address the effect of vector quantization on the computational speed and the accuracy of the reported metrics. Furthermore, as the number of power injection profiles increases, so does the number of ‘scenarios’ to be computed, which impacts the computational speed. In the following chapter, the effect of vector quantization on the QSTS simulation is discussed.

4.6.3 Comparison of the VQ algorithm with other fast time-series algorithms

In this dissertation, the vector quantization algorithm showed attractive computational time reduction. To put it into context, the performance of the algorithm is compared to other approach that were developed as part of an on-going project with Sandia National Laboratories: causal variable time-step [50], predetermined time-step [7], machine learning, event-based [41], [51], intelligent sampling, parallelization [36], circuit reduction [39]. While most approaches showed at least a 2X speed improvement compared to a brute force approach, the vector quantization and the event-based algorithms showed the most promising performances (Figure 20). Furthermore, the scalability of vector quantization algorithm is discussed in this dissertation to insure its applicability.

Hybrid fast time-series algorithms have also been investigated to further reduce the computational speed provided by each individual algorithms [52]. The causal variable time-step algorithm steps through time with an increment equal to the smallest controller delays and if a controller action is experienced, the algorithm backtracks to solve the power flow problem for each time-step [50]. Although capable to adequately approximate the

number of controller action, the computational time reduction is limited by the smallest controller delay – 15 seconds in the modified IEEE 13 bus test case. By combining it with the vector quantization algorithm presented in this dissertation, a hybrid fast time-series algorithm can provide a computational speed improvement by discretizing the load and PV profile and storing previously experienced power flow solutions [52]. On the other hand, the causal variable time-step algorithm can provide a marginal computational speed improvement to the vector quantization algorithm by stepping through time at a large time-step. More specifically, the computational speed improvement comes from avoiding computing some combinations of load and PV multipliers. The benefit of combining the two algorithms is more notable when the profiles are quantized at a finer resolution (larger number of discrete values). Details of the algorithm can be found in [52].

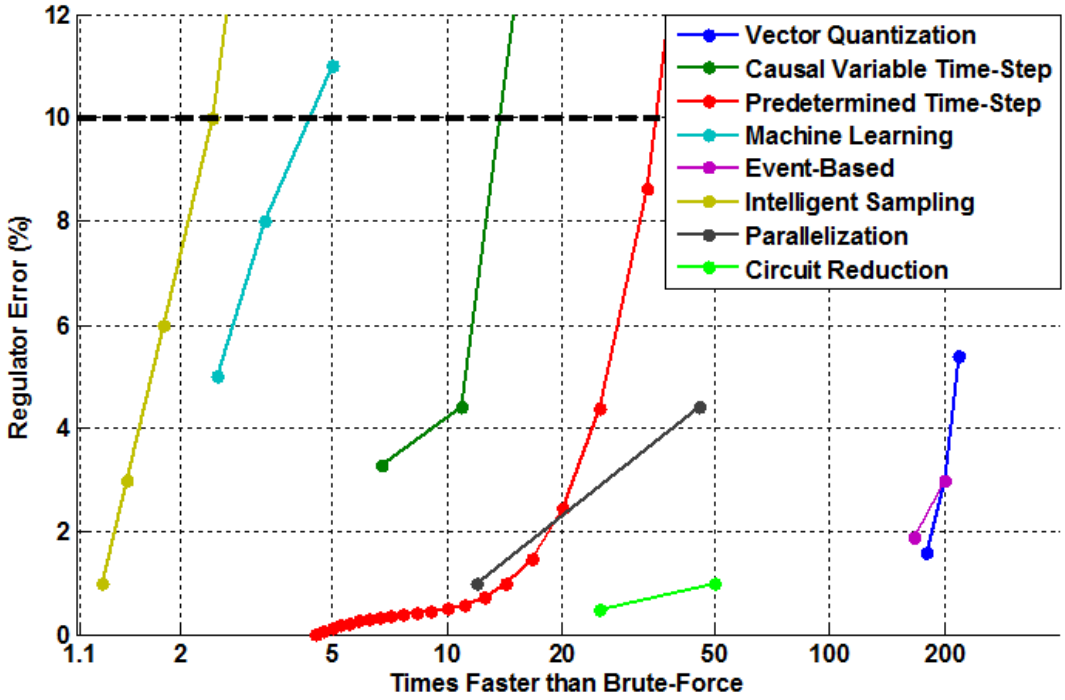


Figure 20 – Performance of various fast time-series algorithms on the modified IEEE 13 bus test case.

CHAPTER 5. EFFECTS OF VECTOR QUANTIZATION ON QSTS METRIC ACCURACY

5.1 Introduction

The speed of the algorithm presented in Chapter 3 and 4 is directly correlated to the number of power flows computed over the time horizon. Through vector quantization, the number of computed power flows can be reduced to decrease the computational time of the QSTS simulation. Vector quantization introduces an error in the profiles to cluster power flow solutions together and, thus, to increase the number of repeated solutions. However, the introduced error can have a significant impact on the accuracy of various metrics reported by the QSTS simulation. The current algorithm does not address how the vector quantization is performed to ensure accurate QSTS results and is specified arbitrarily. Thus far in this work, the performance of the algorithm is determined by running different vector quantization levels and reporting the accuracy of the metrics with respect to the brute-force approach. However, the objective of a fast time-series approximation is to run the QSTS simulation on a feeder and report accurate results with a high confidence level without a brute force simulation. In this chapter, the effect of vector quantization on QSTS metric accuracy is discussed and an efficient method to quantize the profiles is demonstrated.

One of the challenges with correlating the quantization error with the accuracy of QSTS metrics is the complex interactions between controllers in the system [30]. If the error introduced by vector quantization to cluster magnitudes together occurs on the

boundary of a controller action, a small error in the voltage that would otherwise be irrelevant could omit an action from the controller (Figure 21).

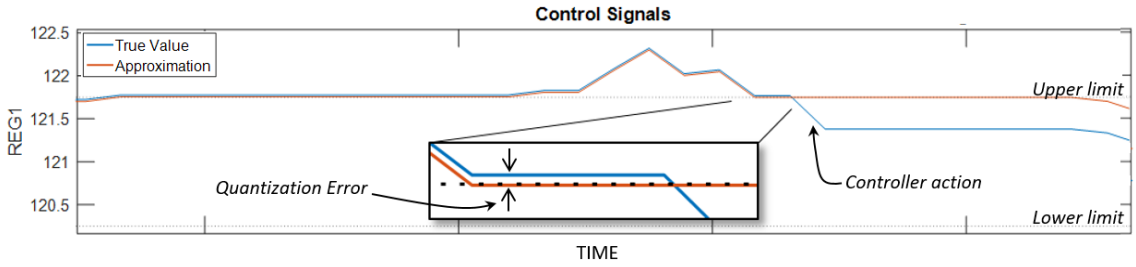


Figure 21 – Control voltage input signal for a regulating tap changer showing how a negligible error can impact on the operation.

Because of incorporated deadbands, this error in the controller state can take a few simulation hours/days to clear. The operation of other controllers can then be drastically altered. For example, a capacitor downstream may have a significant impact on the operation of a load tap changer upstream on the feeder (Figure 22).

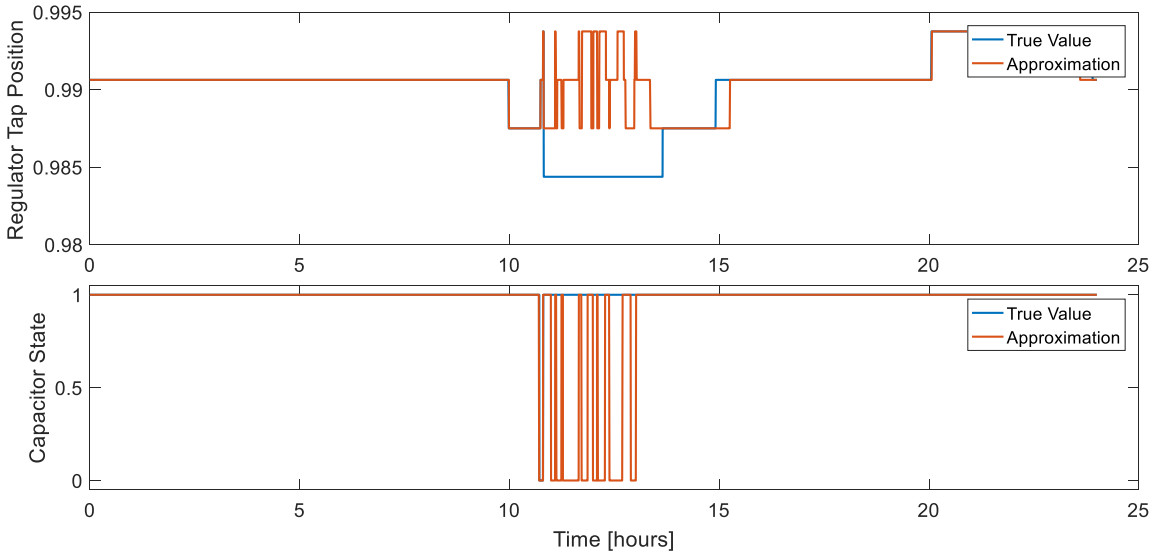


Figure 22 – States of voltage regulating devices over a 24 hour period demonstrating how the capacitor state can create excess actions by the voltage regulator [30].

5.1.1 Random effects of vector quantization on metric accuracies

Quantizing the profiles at different resolutions have a direct impact on the accuracy of the metric accuracy as demonstrated in Chapter 3. Simulation with profiles with higher resolutions should yield better metric accuracy since the introduced error is not as significant. However, in certain cases, it may be possible that some simulations with lower resolution (fewer number of clusters) in a profile yield better accuracy than one with higher resolution (larger number of clusters). This is due to random circumstances where the error introduced by the quantization causes or misses a controller action. If that specific situation is repeated multiple times over the time horizon, it can cause significant errors in the different reported metrics. This could occur at any quantization levels but simulations with profiles quantized at higher resolution are less likely to experience it since the probability that the quantization error causes a controller action is lower. The randomness of this effect can be explored and reduced with multiple simulations at the same quantization level (same number of clusters). By slightly shifting the clusters to move the introduced error, the resolution in the profile is maintained but the random effect of vector quantization is reduced. An illustrating example of shifting clusters while maintaining the same resolution can be found in Figure 23. Notice that in all three cases, the size of each cluster is identical (0.01 p.u.) but the representing value is different. Moreover, note that the representing values are still in the middle of the clusters.

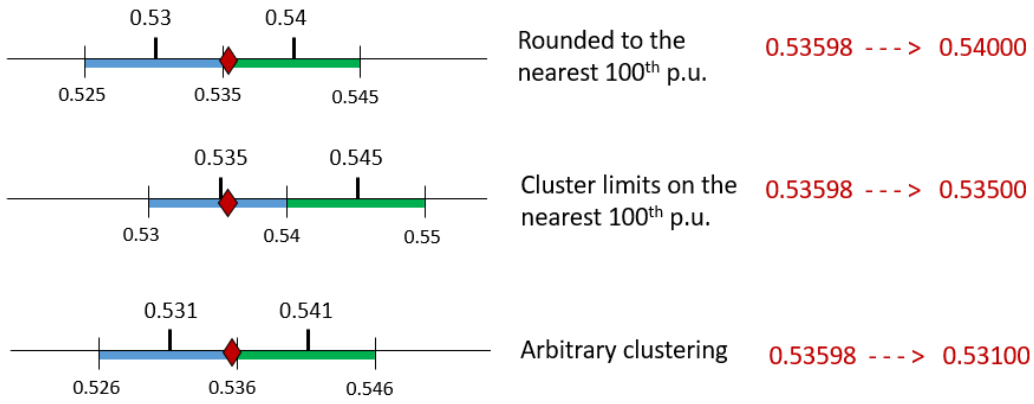


Figure 23 – Illustrative example of shifting clusters while maintaining the cluster size.

In Figure 24, the accuracy of the reported number of tap actions by REG1 in the modified IEEE 13 bus test case is plotted as a function of the shift in the clusters. For instance, a 0.4 coefficient represents a QSTS simulation with the load profile shifted by 40% of the size of the cluster. By carrying this analysis when the load is quantized at 32 and 100 clusters, these simulations results show that there is a higher probability to have a significant error at a lower quantization level.

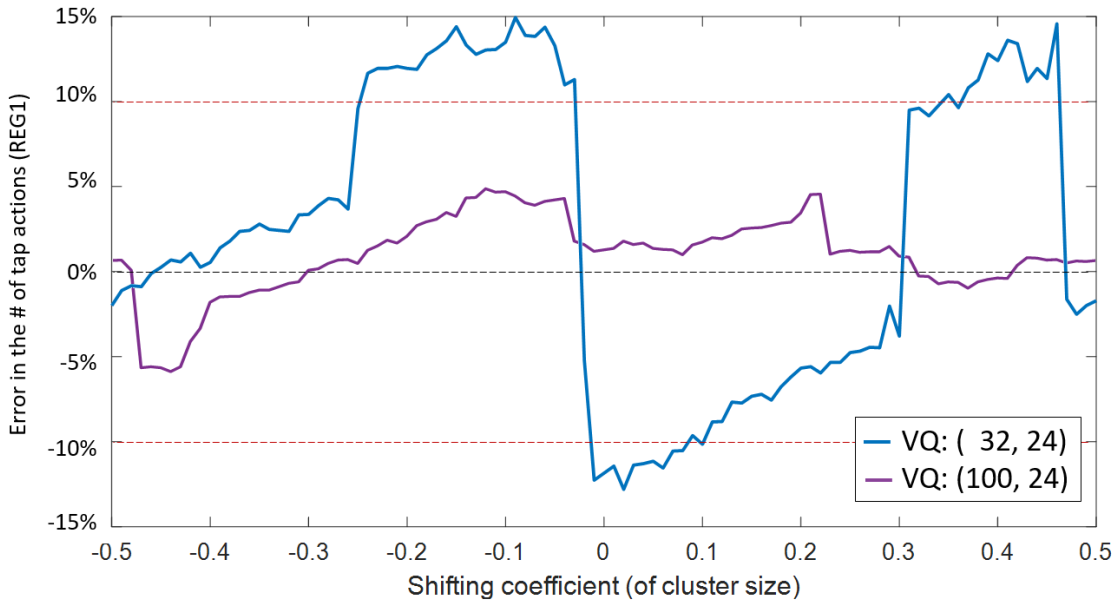


Figure 24 – Error in the number of tap actions for REG1 in the modified IEEE 13 bus test case as a function of the cluster shifts for two different quantization levels (Load, PV).

5.1.2 Background on vector quantization

Given a signal *x*, a vector quantization $Q(x)$ can be obtained by discretizing the signal over its range to reduce the number of discrete amplitude values. This information reduction process introduces an error in the signal by clustering similar amplitudes together. Quantization over the range of data can be done in two ways: a) uniformly, or b) non-uniformly [53]. Uniform quantization, also known as memoryless quantization, clusters amplitudes evenly across the range regardless of the distribution of the data. Non-uniform quantization minimizes the introduced error by properly sizing the clusters and provides additional clusters in denser regions of the range. The introduced quantization error or quantization noise defined as $\epsilon_q = x - Q(x)$ is the difference between the original signal *x* and its quantized signal $Q(x)$. For a discrete-time quantized-amplitude signal, the quantization distortion is the mean square quantization error (MSQE) defined as:

$$MSQE = \sum_{i=1}^{M} \int_{C_i} f_X(x)\ (x_i - Q(x_i))^2\ dx, \tag{17}$$

where C_i is cluster i, $f_X(x)$ is the probability density function such that $x \in C_i$, and M is the number of clusters. The quantizing to assign $Q(x_i)$ can be done uniformly or non-uniformly to minimize the MSQE [54]. In this dissertation, the vector quantization is done uniformly to improve the speed of quantization logic, the accuracy at extreme points, and for higher data compression. This uniformity is especially important when considering the discontinuous nature of the discrete logic of controllable devices on a feeder.

5.2 Characterization of the quantization error (derivations)

The challenge in correlating the error introduced from the quantization process and the accuracy in the various metrics reported by the QSTS simulation is due to the non-linear and more specifically discontinuous nature of the system. The non-linearity in the system comes from the power flow equations and the discontinuity comes from the controllable elements affecting the voltages on the system [41]. In this section, a characterization of the vector quantization error is presented for feeders with and without voltage-regulating equipment.

5.2.1 Feeders without voltage-regulating devices

Without controllers introducing discontinuities in the system, the quantization error can be correlated to the error in the power flow solution. Let us consider the variation of the voltage at any bus *i* on a feeder without voltage regulating devices as a function of the power injection of a PV system (Figure 25). As the output of the PV system increases, the local voltage will also increase creating a continuous function. Furthermore, let us assume that the variation in the power injection is nominal and the distribution feeder operates closely to nominal voltage (i.e. 1 p.u.); thus, the voltage is fairly linearly correlated to the power injection [41]. Through a sensitivity analysis, the slope of the voltage plot can be approximated from the variation in the voltage magnitude at bus *i* (δV_i) and the change in PV output power (δP_{pv}). Obviously, this slope will be different depending on the location of bus *i* and the PV system in the feeder.

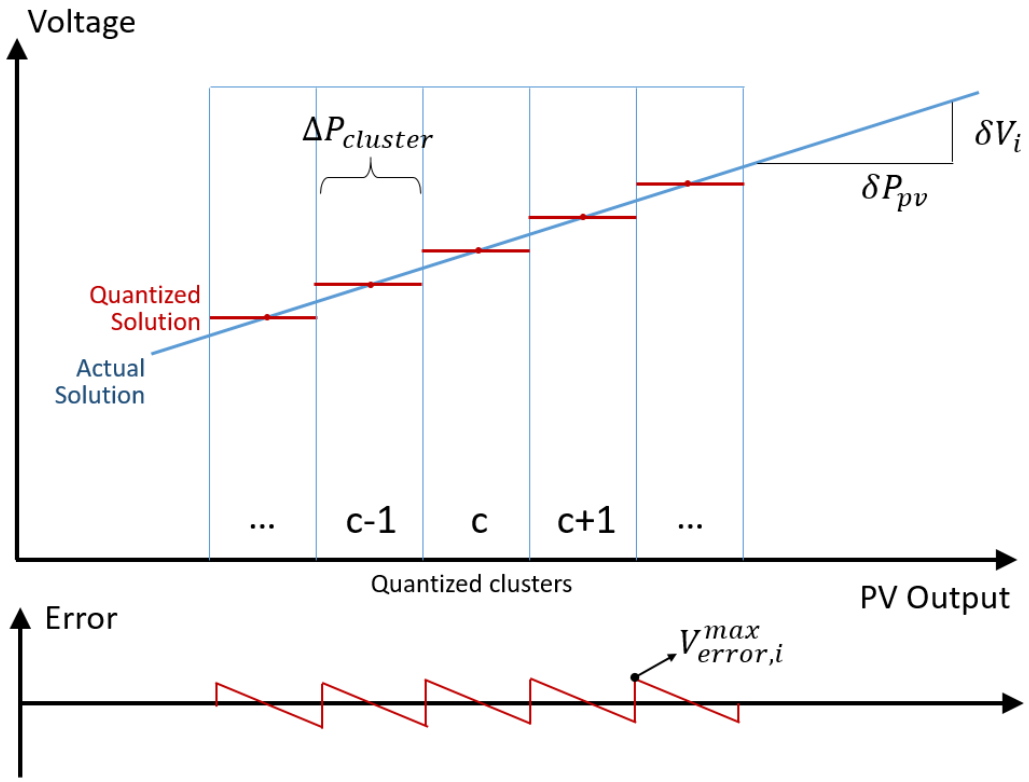


Figure 25 – Approximation of the voltage as a function of the PV output at bus *i* on a feeder without voltage regulating devices.

By uniformly quantizing the PV profile (see Figure 25), it creates discrete jumps in the output power of the PV systems, $\Delta P_{cluster}$, which can be calculated with the following equation:

$$\Delta P_{cluster} = \frac{P_{pv,max}}{\left(N_{cluster,PV} - 1\right)} \tag{18}$$

where $P_{pv,max}$ is the maximum power output of the PV system, and $N_{cluster,PV}$ is the number of clusters (i.e. discrete values) over the range of PV values. As a result, the power flow solution (e.g. nodal voltage) also sees discrete jumps. Thus, the maximum absolute error in the voltage magnitude at bus *i* can be characterized by:

$$V_{error,i}^{max} = \frac{\Delta P_{cluster}}{2} \frac{\delta V_i}{\delta P_{pv}} \tag{19}$$

Assuming a uniform distribution of the data points across each cluster, the mean absolute error due to quantization is defined as:

$$V_{error,i}^{mean} = \frac{V_{error,i}^{max}}{2} \tag{20}$$

If the objective of the QSTS simulation is to estimate the voltage extremes experienced throughout the time-horizon, the maximum error due to quantization defined in Eq. (19) is the upper bound of the error in the metric. Furthermore, the maximum error in other measurements on the feeder (e.g. current, power losses, etc.) can be similarly determined from a linear sensitivity analysis.

5.2.2 Feeders with voltage-regulating devices (discontinuities)

A vast majority of distribution feeders have voltage-regulating devices to improve power quality on the system. However, these voltage-regulating equipment (e.g. load tap changers or capacitor banks) create discontinuities in the nodal voltages on the feeder [41] because of the discrete states of the devices – on/off state of capacitor banks or the tap position for LTCs. As soon as a controller changes state, it creates a discrete shift in the voltages on the feeder. The challenge associated with these discontinuities is that the quantization error can occur near these discrete voltage regulations and create a significant error in the power flow solution that would not be otherwise be present in the profiles (see Figure 26). Some measurements are more affected by these discontinuities than others, namely, the voltage magnitude. Thus, this section focuses on the error in the voltage magnitude.

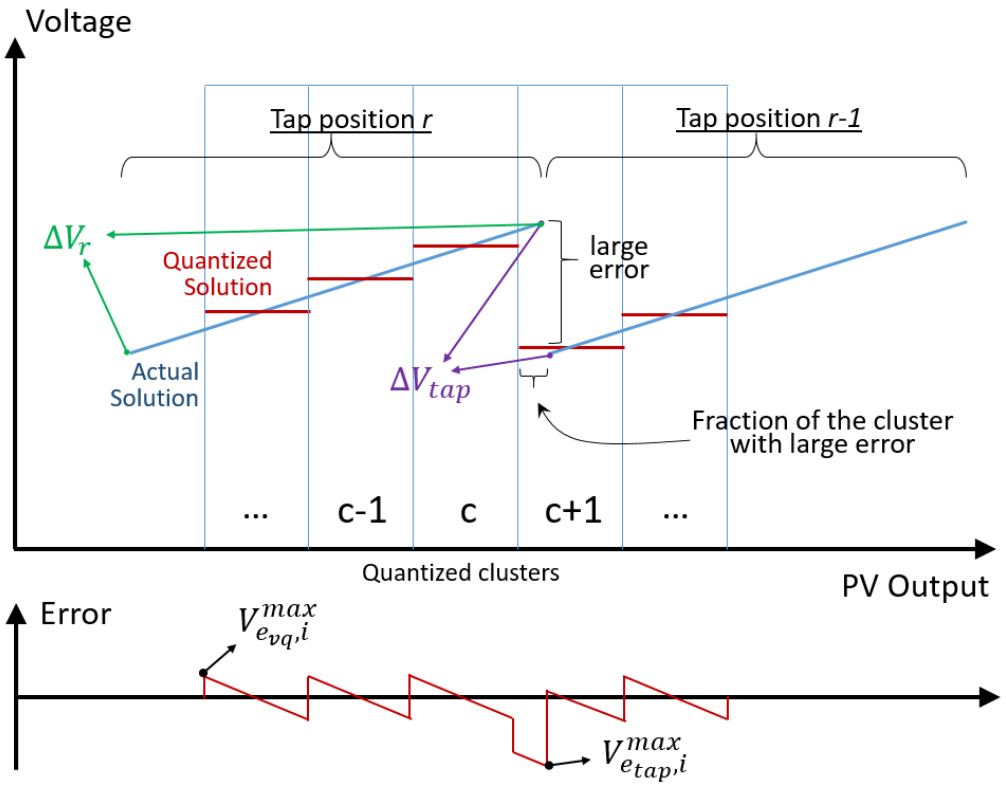


Figure 26 – Approximation of the voltage as a function of the PV output at bus *i* on a feeder with voltage regulating tap changer.

Although the quantization error on the profile remains the same whether or not the feeder has voltage-regulating equipment (and hence discontinuities), the error in the voltage is more significant around the decision boundaries of the controllers where an action occurs. Ideally, the quantization could be performed to occur exactly where a controller action would be performed to avoid this large error. However, because some feeders have multiple voltage-regulating devices and multiple power injection profiles, it become extremely complex to quantize the profiles in such a way that the clusters do not overlay on top of a controller action. Furthermore, deadbands in the controllers add another layer of complexity where a specific power injection scenario can provide multiple feasible solutions due to the controllers (see Challenge 4 in [30]). The effect of deadbands and delays on the error is discussed in section 5.2.3.

The percentage of data points with a large error can be analytically studied. Let us consider the change in PV output power only (constant load) and a single voltage-regulating device (i.e. a regulator) without deadbands. The change in PV output power (ΔP_r) required to change one tap position from *r* to *r*-1 is equal to:

$$\Delta P_r = \frac{\Delta V_r}{\frac{\partial V_i}{\partial P_{pv}}} \tag{21}$$

where ΔV_r is the potential voltage range for the tap position before the controller would change the tap up or down, and $\frac{\delta V_i}{\delta P_{pv}}$ is the rate of voltage change with respect to the power variation. Let us also define the shift in voltage created by one tap change as ΔV_{tap}. To avoid oscillation, ΔV_r must be equal or larger than ΔV_{tap}, else the shift in voltage created by one tap change could keep oscillating outside the upper and lower voltage thresholds. In this section, ΔV_r is equal to ΔV_{tap} under the assumption that the voltage regulator does not have a deadband. This assumption is relaxed in the following section.

Because of the discontinuities, the maximum absolute error in the voltage due to the quantization of the PV output multiplier can be larger than the one defined Eq. (19). The upper bound of the maximum absolute error is experienced when the tap action occurs exactly in the middle of a cluster and is equal to the shift in voltage created by a tap action:

$$V_{error,i}^{max} = V_{e_{tap},i}^{\max} = \Delta V_{tap} \tag{22}$$

As shown in Figure 26, only a fraction of the PV output power multipliers would have a large error depending on how the multipliers are quantized. The number of clusters included over the change of PV output power that would trigger a tap change is equal to:

$$N_{cluster,PV,tap} = \frac{\Delta P_r}{\Delta P_{cluster}} \tag{23}$$

Let us consider that $N_{cluster,PV,tap}$ is an integer and is equal to 10. If the rightmost cluster (*c+1*) is on the boundary, a fraction of the cluster will have a significant error in the voltage due to the different controller state. Since the representing values for each cluster is at its center, the fraction of the cluster that would have a significant error is at most ½ of the cluster size or $\frac{\Delta P_{cluster}}{2}$. Under the assumption that all values are evenly distributed in each cluster, the percentage of data point with a wrong controller state would be at most equal to $\varepsilon_{large\ error} = \frac{1}{2N_{cluster,PV,tap}}$ (or 5% in this example). Thus, the upper bound for the mean absolute error can be defined in terms of the error due to the vector quantization ($V_{e_{vq},i}^{max} = \frac{\Delta P_{cluster}}{2} \frac{\delta V_i}{\delta P_{pv}}$) and the error due to the change in a tap position ($V_{e_{tap},i}^{max} = \Delta V_{tap}$):

$$\begin{aligned} V_{error,i}^{mean} &= \left(1 - \frac{1}{2N_{cluster,PV,tap}}\right)\left(\frac{V_{e_{vq},i}^{max}}{2}\right) + \frac{1}{2N_{cluster,PV,tap}}\left(\mathrm{V}_{\mathrm{e_{tap}},i}^{max} - \frac{V_{e_{vq},i}^{max}}{2}\right) \\ &= \left(1 - \frac{1}{N_{cluster,PV,tap}}\right)\left(\frac{V_{e_{vq},i}^{max}}{2}\right) + \frac{1}{2N_{cluster,PV,tap}}\Delta V_{tap} \end{aligned} \tag{24}$$

By increasing the number of clusters between tap changes, the mean absolute error can be reduced and ultimately converges to $V_{e_{vq},i}^{max}$.

This derivation is leveraged in a vector quantization method (also referred as strategy) to determine an appropriate quantization level for each profile. Let us assume that one would want the percentage of data points with a significant error (i.e. due to a tap change) to be below a certain percentage, $\varepsilon_{large\ error}$. The number of clusters over the

range of PV output power multipliers could be determined from Eq. (18), (21) & (23) with the following equation:

$$N_{cluster,PV} = \left(\frac{1}{2\,\varepsilon_{large\ error}}\right)\left(P_{pv,max}\frac{\partial V_i}{\partial P_{pv}}\right)\left(\frac{1}{\Delta V_r}\right) + 1 \tag{25}$$

where $\left(P_{pv,max}\frac{\partial V_i}{\partial P_{pv}}\right)$ is the variation in the voltage at bus *i* created by the profile if it were to change from 0 to $P_{pv,max}$, and normalizing this voltage variation with ΔV_r determines how many controller actions would have occurred over that change. For example, if the controller would have seen 5 actions over that range and if the objective is to have at most 5% of data points with a large error, the profile should be quantized to have 51 discrete values over its range.

The maximum and mean absolute error derived in this section are applicable when analyzing the time-series data from the simulation. However, if the objective of running a QSTS simulation is to capture the voltage extremes, the error will remain within $V_{e_{vq},i}^{max}$ since the large error discussed above occurs when regulating the voltage. On the other hand, if the objective of the QSTS simulation is to study the operation of voltage regulating devices (i.e. the number of controller actions), the error in the number of discrete actions cannot be correlated to the quantization error in the profile without running through the time-series simulation. Moreover, deadbands and delays add an additional layer of complexity.

5.2.3 Effects of overlapping deadbands, delays, and controller interactions

Although the upper-bound for the percentage of time-steps in a yearlong simulation with a large error is derived above, this is under the assumption that $\Delta V_r = \Delta V_{tap}$ and that the controllers do not have integrated delays (e.g. 15 seconds before taking an action). The impact of overlapping deadbands in the derivation is that the number of time-steps with wrong controller states is dependent on which state it is in. As discussed in [30], there can be multiple valid controller states within their deadbands for a specific power injection. This is because the regulation in the voltage by one discrete controller action is often smaller than its deadband. For example, the LTC in the modified IEEE 13 bus test case has 33 taps over $0.1\text{V}_{\text{p.u.}}$ (or $12V_{120V\,base}$) making each tap change a $\Delta V_{tap} = 0.375\text{V}_{120\text{V}\,base}$ while its deadband is $\Delta V_r = 1.5V_{120V\,base}$. Thus, $\Delta V_r = 4\Delta V_{tap}$ in this test case. Again, the objective of having a deadband larger than the voltage regulation is to reduce oscillations in its operation. An illustrative example of overlapping deadbands can be found in Figure 27.

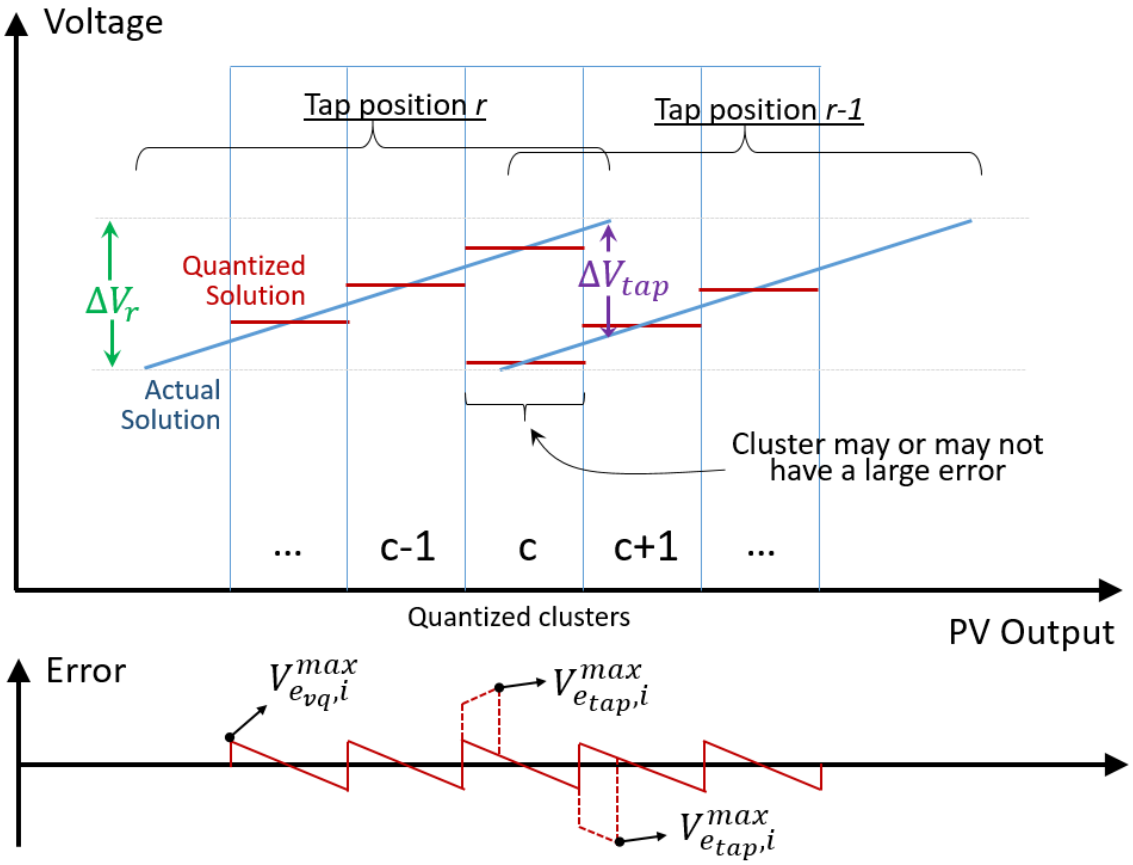


Figure 27 – Illustrative example of the effect of a deadband in voltage-regulating devices on the error created by quantization.

As the PV output power increases on a circuit with a regulator on tap position *r*, it may reach a voltage threshold that will trigger a tap action and thus a shift in the voltage on the feeder. After which, the PV output power may decrease until the voltage on the feeder reaches a lower bound threshold that would trigger the regulator to go back to its original state. Because the deadbands are overlapping, the two recorded actions did not occur at the same PV output level. When the PV output is quantized, some clusters of PV output values may or may not have a significant error. Therefore, the percentage of time-steps with a large error is lower than $\varepsilon_{large\ error}$ but cannot be estimated without going through the time-series simulation.

Furthermore, the delays in controllers add another complexity to the problem because a time-step with a large error may not create an additional controller action if the delays never expired before the voltage returns within the deadband. The effect of delays cannot be derived and should be estimated only by stepping through the simulation.

Lastly, controllers also impact one-another where an action from on controller may trigger an action from another controller downstream or upstream because of the voltage shift it created. One example of such interaction is in the modified IEEE 13 bus test case described earlier. The capacitor bank (600kVAR) provides a significant variation in the voltage on the feeder when it changes states that the regulators upstream also change states in response to it. Again, this creates another challenge to correlate the error introduced by the quantization of the profile with the number of controller actions recorded over the time horizon.

As discussed in this sub-section, overlapping deadbands, delays, and controller interactions creates an inevitable hysteresis in the system that is very difficult to study without running a QSTS simulation. However, the voltage sensitivity analysis used to characterize the error for feeders with controllers without deadbands can be leveraged to quantize the profiles efficiently. More specifically, this sensitivity analysis can be used to understand the impact of various power injection profiles (e.g. load or PV) on the different controllers. A vector quantization strategy based on the analysis of this section is proposed in a section 5.4.

5.3 Simple vector quantization strategies

In the previous two chapters, the vector quantization algorithm required the quantization level of each profile to be arbitrarily chosen prior to running the simulation. However, without prior knowledge, the accuracy of the results cannot be estimated without running multiple simulations until the results converge or without running the QSTS simulation with a brute force approach. In both cases, the computational time reduction is negatively impacted. Thus, a method (strategy) to quantize the different profiles is needed to optimize the accuracy of the simulation results for a simulation speed. In this section, two simple vector quantization (VQ) strategies are discussed.

5.3.1 Uniform VQ strategy

The simplest VQ strategy approach is to quantize all the profiles equally to have the same number of discrete values regardless of their impact on the feeder. For example, this approach is used in Chapter 3 on the modified IEEE 13 bus test case simulations by quantizing the load and PV profiles in 101 clusters each. This strategy would be used if no information is available about the specific feeder model. Although it may not yield the most accurate simulation results for a given speed, this VQ strategy is discussed in this dissertation as a reference to other strategies.

5.3.2 kW-based VQ strategy

A more intelligent VQ strategy is to quantize the different profiles based on the kW value associated with it. For instance, in the modified IEEE 13 bus test case, the profiles for the load and the PV system are reported in per unit of nominal power but the PV system

is sized at ~40% of peak load. Thus, if the profiles are both quantized to create 101 clusters (0.01 p.u. resolution), an incremental jump of 0.01p.u. in each profile would not create the same change in net power injection change on the feeder. A VQ strategy can be based on the kW value by simply quantizing the profile respectfully and an incremental jump from one discrete value to another would be equivalent to the same net power injection change in each profile. For instance, if the load profile is quantized to create 101 clusters (or rounded to the nearest 0.01 p.u.), the PV profile would be quantized to create 41 clusters (rounded to the nearest 0.025 p.u.). Performance results of this strategy can be found in section 5.5.

5.4 Proposed VQ strategy

The number of controller actions is arguably the most challenging metrics to capture from the QSTS simulations because of the discontinuous nature of the controller and their complex interactions. Each profile will affect the various controllers differently based on their kW value as well as the location of their power injections. While the kW-based VQ strategy quantizes the profiles based on their power injections, one crucial aspect that is not included with this approach is the consideration of the location of the power injection. For instance, the location of a PV system can impact the operation of a voltage regulator differently whether it is located closely downstream or at the end of the feeder. Thus, the proposed VQ strategy is to perform a sensitivity analysis of the impact that each profile has on each controller, which would consider the location of the power injection. This is done by introducing a small disturbance in each profile individually and recording the controller input signals. This sensitivity analysis allows to characterize the variation in each controller input signals that a profile creates. In other words, this analysis provides an

understanding of the variation in the controller input signals that is created when a profile jumps from one cluster (discrete value) to another.

To conduct this sensitivity analysis on a new test circuit, each profile is first set to their mean value ($\bar{X}_n$). The power flow equations are then solved with the controllers enabled to determine feasible states within their deadbands. The controllers are then disabled to avoid any variations in the power flow solution due to a change of state. For each profile, two power flow solutions are computed when a small disturbance (δ_n) is introduced into the profiles individually and the input voltage signal to each controller *c* ($V_{n,c}$) is recorded to compute the sensitivity. The change in the input voltage signal is then normalized with respect to the disturbance. The algorithm of this sensitivity analysis is illustrated in Figure 28.

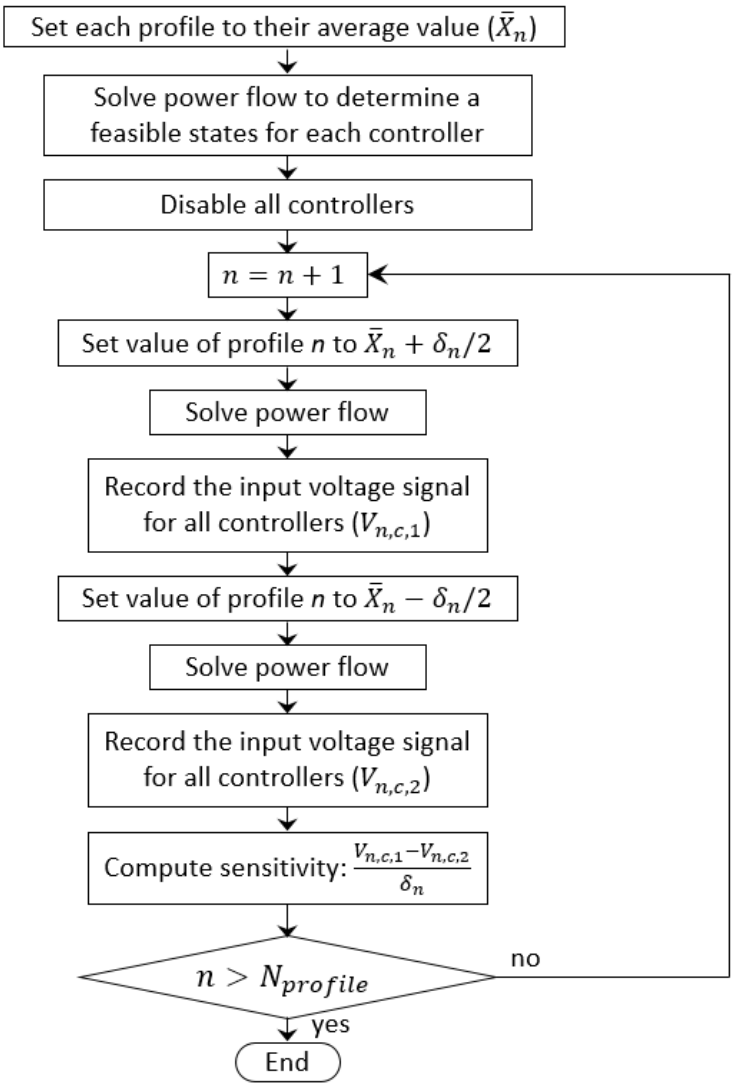


Figure 28 – Flow chart of the sensitivity analysis algorithm conducted for the proposed VQ strategy.

The resulting voltage sensitivity represents the input control voltage variation that a 0 to 1 p.u. change in the profile would create in each controller. This value can then be normalized with respect to the controller deadbands (ΔV_r) to determine how sensitive a controller is to a specific profile. Each profile can then be quantized according to the highest impact it has on any controllers. Thus, the vector quantization strategy as part of the QSTS simulation can be implemented prior to the time-series simulation to quantize the profiles as shown in the flow chart in Figure 29.

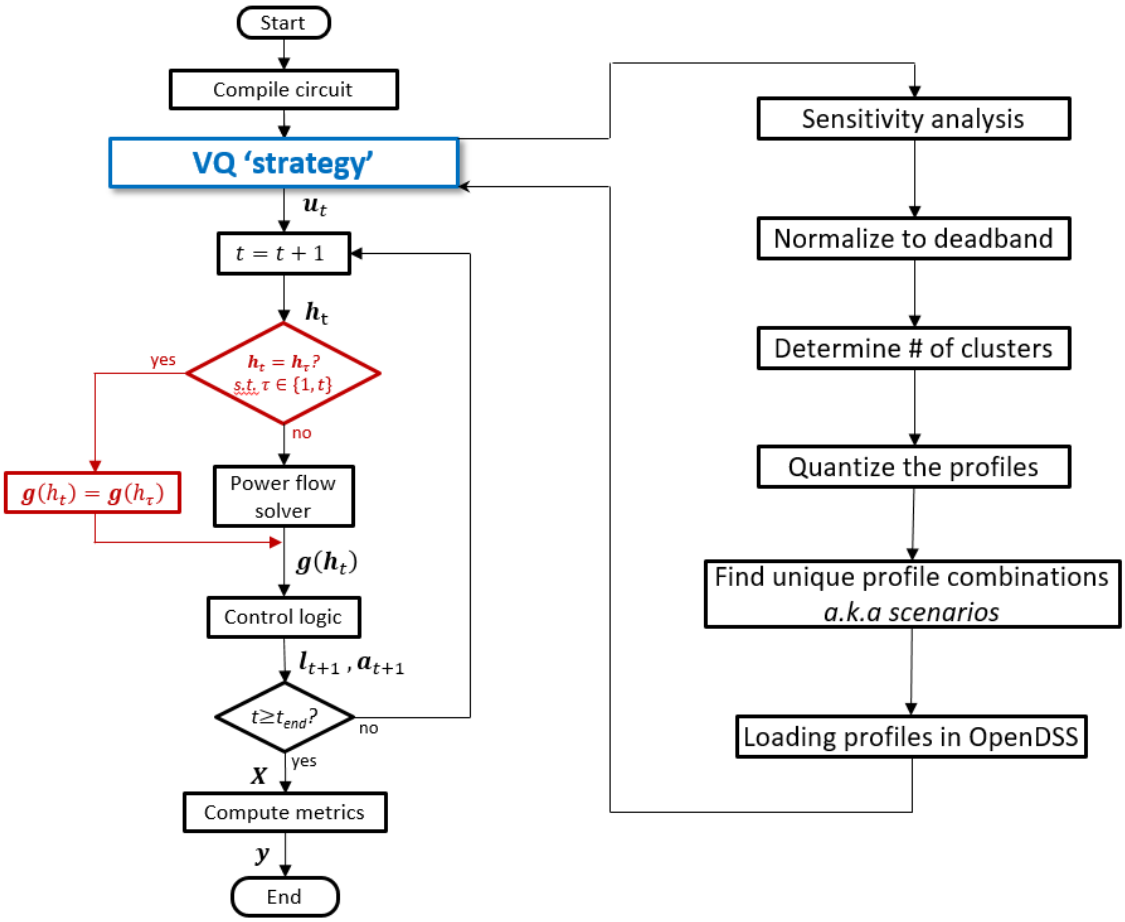


Figure 29 – Flow chart of the proposed VQ strategy based on voltage sensitivity analysis.

This sensitivity analysis is carried out on the modified IEEE 13 bus test case. In TABLE 7, the sensitivity analysis is reported as a percentage of the deadband that a 0 to 1.0 p.u. change in each profile would create. For instance, if the PV output power were to change from 0 to 100% of its nominal power, the capacitor bank would see a change in its controller input signal of 113% (of the deadband) – or $3.39V_{120base}$. Note that the load creates a negative variation in the controller input signals and the PV system creates a positive variation because of the direction of the power flow injection affecting the voltages.

TABLE 7. Sensitivity analysis of the impact of each profile (1p.u. variation) on the different controllers in the modified IEEE 13 bus test case as a percentage of the deadband.

	SUB-a	SUB-b	SUB-c	CAP1
Deadband	1.5 V	1.5 V	1.5 V	3.0 V
Load	-438 %	-438 %	**-565 %**	-352 %
PV	80 %	82 %	83 %	**113 %**

Obviously, the profiles would need to be quantized in such a way that the difference between each discrete value is less than 100% of the deadband. For example, if the profiles were quantized to create 11 clusters each, the variation in the controller signal created by a cluster 'jump' would be $1/10^{th}$ of the values in the table above. The proposed approach is to quantize the profiles based on the ratio of their highest impact (5.0:1 ratio in this case). Furthermore, they can be quantized in such a way that the highest variation created by that profile in any controller input signals would be below a pre-determined percent threshold for all controllers. For instance, the load and PV profiles would need to be quantized into 57 and 12 clusters, respectively, to maintain an absolute change from one cluster to another below 10% of the respective deadbands (see TABLE 8). In other words, if either the load or PV output were to change from one discrete value to another, the largest variation in any controllers would be -9.9%, which is below the targeted 10%.

TABLE 8. Impact of a cluster 'jump' in each profile on the different controllers in the modified IEEE 13 bus test case.

	SUB-a	SUB-b	SUB-c	CAP1
Load (57 clusters)	-7.7 %	-7.7 %	-9.9 %	-6.2 %
PV (12 clusters)	6.7 %	6.8 %	7.0 %	9.4 %

By varying the pre-determined percent threshold, a curve of the computational speed versus accuracy can determine whether or not the proposed VQ 'strategy' is optimal. This is tested on the modified IEEE 13 bus test case (see section 3.5 for details) and the

large distribution system feeder test case (see section 4.3 for details) in the following simulation section.

5.5 Simulation Results

5.5.1 Modified IEEE 13 bus circuit

Before testing the performance of the different VQ strategies, the effect of vector quantization on the accuracy of the QSTS simulation is tested on the modified IEEE 13 bus test circuit by quantizing the load and PV profiles at different levels and recording the number of controller actions of the load tap changer on phase a (REG1) for each simulation. A heat map of the absolute percent error is plotted in Figure 30 and three traces highlight simulations with a specific load to PV profile quantization ratio. The 1:1 ratio represents the uniform VQ strategy while the 5:2 and 5:1 ratio represents the kW-based and the proposed VQ strategy, respectively. In addition, the computational time reduction for each simulation is reported in Figure 31.

The results in the figures below show that the load profile has a more significant impact than the PV profile on the accuracy of the reported metric. Note that the three marked simulations (Point A, B, & C) have the same computational speed but reported different metric accuracies. Thus, quantizing each profile equally may not optimize speed and accuracy and both the proposed VQ strategy and the kW-based method are better than the uniform VQ strategy.

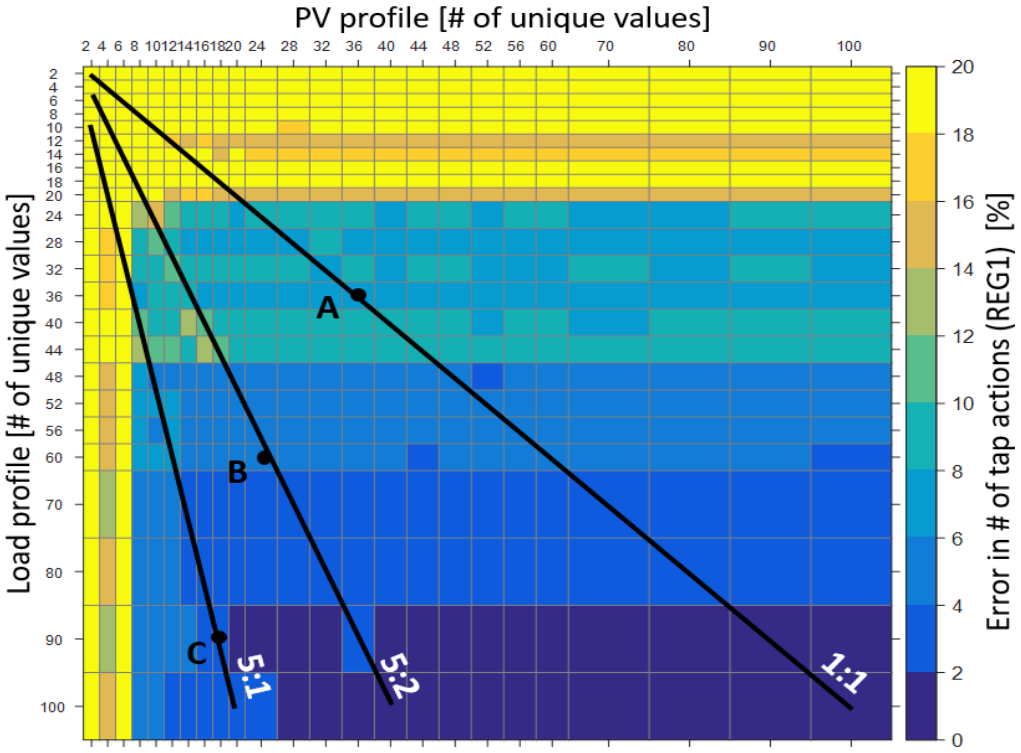


Figure 30 – Percent error of the number of tap actions (REG1) reported by QSTS simulations as the profiles are quantized at different levels (x- and y-axis represent the number of clusters for each profile).

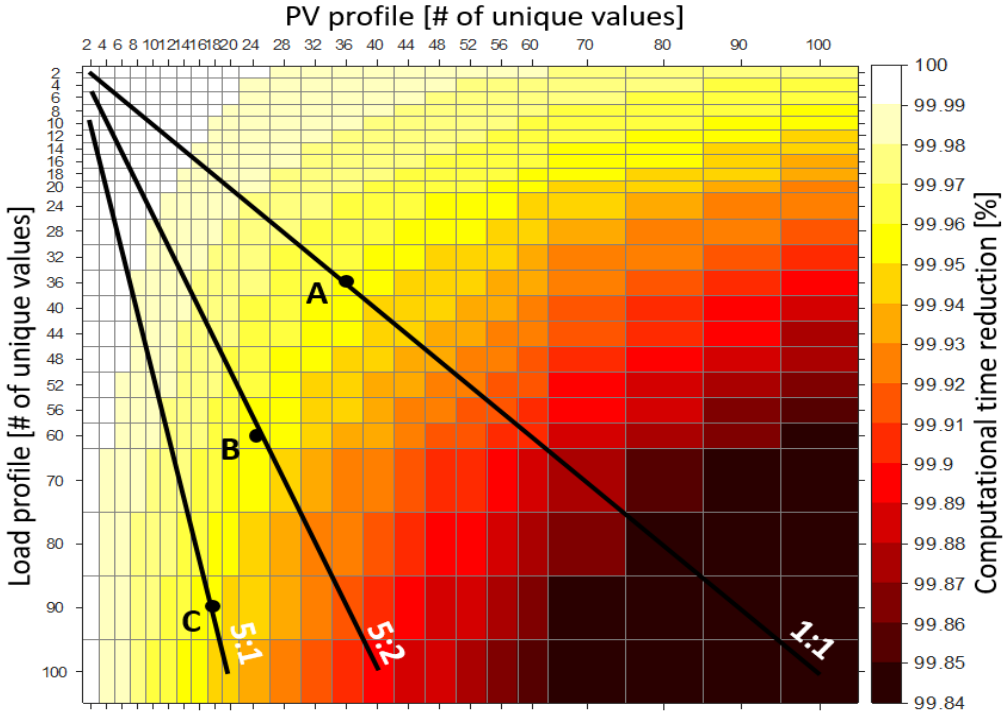


Figure 31 – Heat map of the computational time reduction due to the vector quantization algorithm.

Another way of demonstrating that the proposed strategy is a reasonable approach to quantize the profiles is by plotting the error versus the computational speed. In Figure 32, the RMS error of the number of tap actions recorded by all regulators is plotted as a function of the computational time reduction for all QSTS simulations in the heat maps above. Traces are also included to highlight the simulations conducted under the three vector quantization strategies.

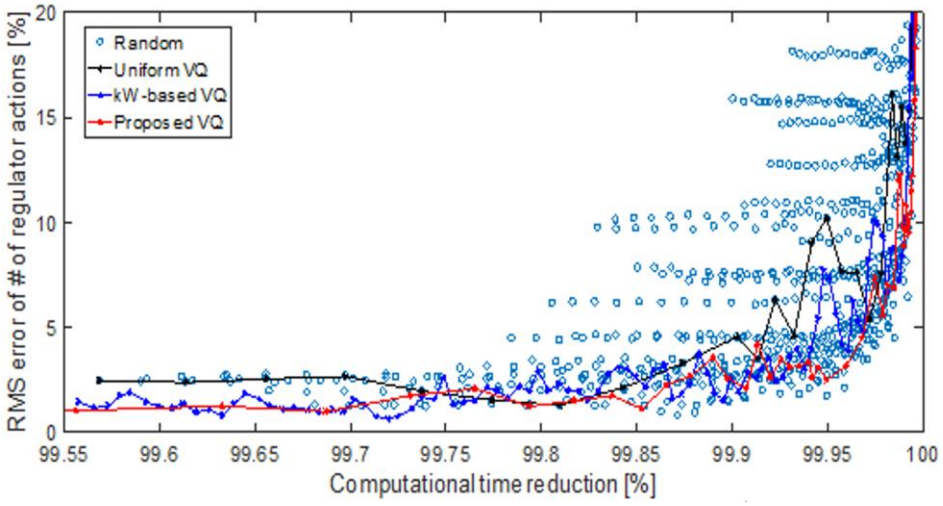


Figure 32 – Speed versus accuracy of the 576 simulations in the heat maps above.

As expected, the kW-based VQ strategy and the proposed VQ strategy are closely related and both are close to the pareto front. Quantizing the profiles according to one of these methods should yield more accurate results than a simulation with both profiles quantized equally. Although the difference in the performance between the proposed sensitivity-based strategy over the kW-based approach is not clear in this test case, the advantages of the proposed method becomes more apparent for test cases with more profiles and where the location of their power injections impacts controllers differently (see the large distribution feeder test case in the following subsection).

The reason some random simulations are below the trace of the proposed strategy is due to the interactions between controllers where a controller can significantly impact the operation of another regulators [30]. In this modified IEEE 13 bus test case, the controllable capacitor bank has an impact on the operation of the regulators because of the reactive power it injects. Because controller interactions cannot be easily characterized without running through a time-series simulation, the proposed vector quantization strategy provides an attractive compromise between speed and accuracy.

5.5.2 Large distribution feeder

A similar approach is taken on the large distribution feeder test case where a sensitivity analysis is first carried out. In the table below, the variation in the controller input signals created by a 0 to 1p.u. change in the profiles is normalized with respect to the deadband of the controllers.

TABLE 9. Sensitivity analysis of the impact of each profile on the different controllers. Values are reported in percentage of the deadband when the profiles vary by 1 p.u.

	SUB	REG-a	REG-b	REG-C	CAP1	CAP2	CAP3	CAP4-a	CAP4-b
Deadband	**2.0 V**	**2.0 V**	**2.0 V**	**2.0 V**	**3.0 V**	**3.0 V**	**3.0 V**	**5.0 V**	**5.0 V**
Residential (4.2MW)	-117 %	-240 %	**-456 %**	-373 %	-138 %	-242 %	-261 %	-231 %	-380 %
Commercial (13.7MW)	-427 %	-517 %	**-531 %**	-527 %	-332 %	-391 %	-377 %	-291 %	-298 %
PV1 (2.5MW)	46 %	118 %	**121 %**	116 %	61 %	94 %	80 %	55 %	57 %
PV2 (0.5MW)	9 %	26 %	25 %	23 %	14 %	32 %	44 %	**49 %**	46 %
PV3 (0.4MW)	21 %	**50 %**	2 %	-13 %	37 %	33 %	28 %	19 %	1 %
PV4 (0.4MW)	6 %	17 %	-6 %	**34 %**	9 %	17 %	15 %	10 %	0 %

Similarly, the 6 profiles can be quantized in such a way that the largest step in a profile is below a specific percent threshold of the deadbands. This is carried out over a range of thresholds and is traced in the figure below. Because of the number of profiles and size of the circuit, not every combination of quantization levels can be simulated. Thus, random vector quantization levels have been performed on the profiles and the accuracy of the reported metrics have been recorded for more than 250 simulations. In addition, the uniform and kW-based VQ strategies provide a perspective on the efficacy of the proposed VQ strategy.

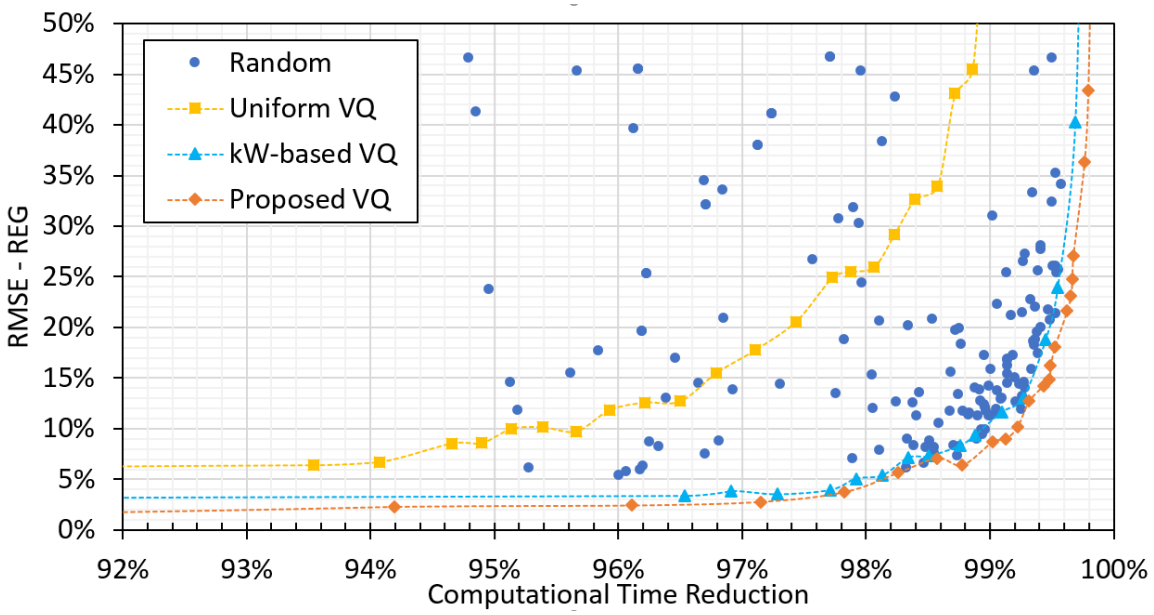


Figure 33 – Computational time reduction versus the root mean square error for all tap changers in the large realistic distribution test feeder.

Again, the proposed VQ strategy based on the sensitivity analysis is at the Pareto front. The kW-based VQ strategy also showed very attractive results as the sensitivity analysis is closely related to the kW value associated with each profile. Although one could argue that 250 QSTS simulations is not representative of all quantization possibilities, this proposed VQ strategy is extremely quick and provides valuable insights to optimize

computational speed and accuracy and performed better than all 250 other simulations. Note that because the circuit is larger and there is more reactive power flow due to the capacitors, the location of the power injection does impact the controllers differently and their profiles should be quantized accordingly.

5.5.3 Other feeders: variations of the previous two test cases

Thus far, the results presented in this section shows that a reasonable approach to quantize the profile is to base it on the variation that each creates in the controller input signals as a percentage of controller deadbands. However, a specific percentage threshold has not been recommended to remain within acceptable errors. In this subsection, a few additional test cases are simulated to study the relationship between the percentage threshold and the controller action errors.

The RMS error of the number of actions of all LTCs can be plotted as a function of the percentage of the deadband to determine an appropriate threshold that would yield acceptable errors. Since this comparison could vary from one feeder to another, four variations to the modified IEEE 13 bus test case are simulated where the PV system size and the regulator deadband size are varied. TABLE 10 summarizes the five test cases.

TABLE 10. Variations of the modified IEEE 13 bus test case

NAME	PV SYSTEM SIZE	DEADBAND (120V BASE)
IEEE13-REFERENCE	2000 kW	1.5 V
IEEE13-1750	**1750 kW**	1.5 V
IEEE13-2500	**2500 kW**	1.5 V
IEEE13-1V	2000 kW	**1.0 V**
IEEE13-2V	2000 kW	**2.0 V**

For each variation of the test case, 26 QSTS simulations are conducted where the profiles are quantized between 0.5% and 25% of the deadband. The RMS error of number of actions from all three regulators are reported in the figure below.

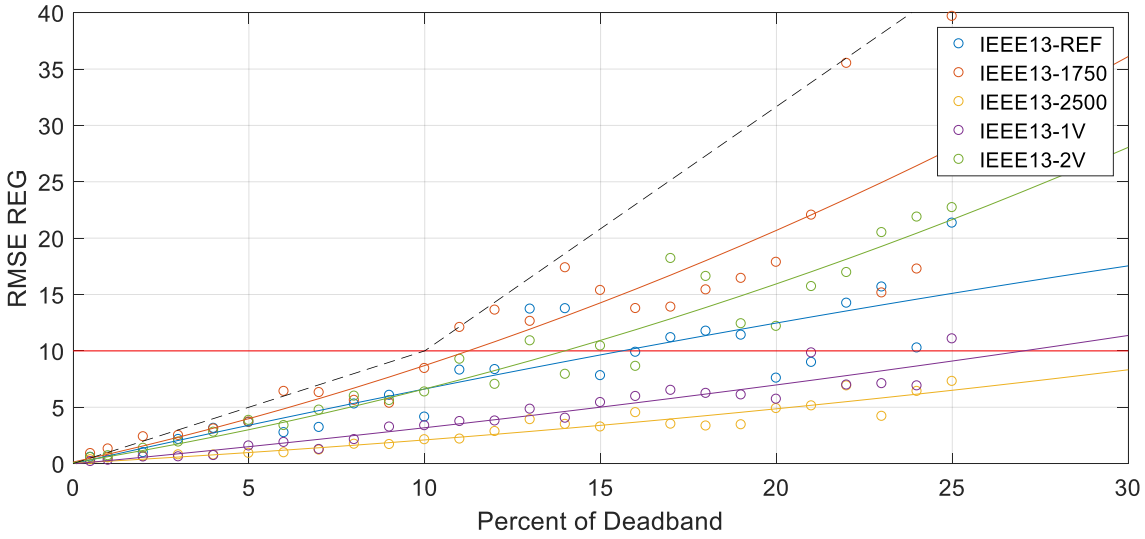


Figure 34 – RMS error of the number of controller action for all regulators versus the percent of deadband used to quantize the profiles for five variations of the IEEE 13 bus test circuit.

For all five variations of the circuit, the RMS errors remain within the 10% error threshold if the profile are quantized with a deadband percentage below 10%. The black trace shows a one-to-one ratio to the RMS error and the percent of the deadband when below 10%. Above that threshold, the error is much more unpredictable because there is a higher probability that the introduced quantization error triggers a controller action. Note that the data cannot easily be traced with trend lines because of the complex controller interactions. Another important observation drawn from these simulations is that the test case variations with the least number of total actions by all controllers (TABLE 11) performed the worst and were more unpredictable at higher thresholds. The reasoning behind this behaviour is that an error in a controller state can propagate for much longer if

the controller did not operate often over the time horizon. This is extensively discussed in section 2.2.5 with supporting figures.

TABLE 11. Number of actions recorded with the brute force approach

NAME	REG1	REG2	REG3	CAP	TOTAL
IEEE13-REFERENCE	7,046	7,220	8,447	2,504	25,217
IEEE13-1750	3,178	3,324	4,439	742	**11,683**
IEEE13-2500	16,862	17,141	19,033	5,698	58,734
IEEE13-1V	16,389	16,915	19,122	798	53,224
IEEE13-2V	4,160	4,331	5,683	4,700	**18,874**

To verify that a 10%-deadband quantization strategy yields errors below 10%, two additional circuits are tested. First, the capacitor bank is turned off in the modified IEEE 13 bus test case to verify that the controller interaction is the source of the error (IEEE13-NoCap trace). Second, the hypothesis is tested on the large distribution feeder test circuit discussed in Chapter 4 (LargeCircuit trace). The RMS error is plotted in Figure 35 where the reference black trace from Figure 34 is included.

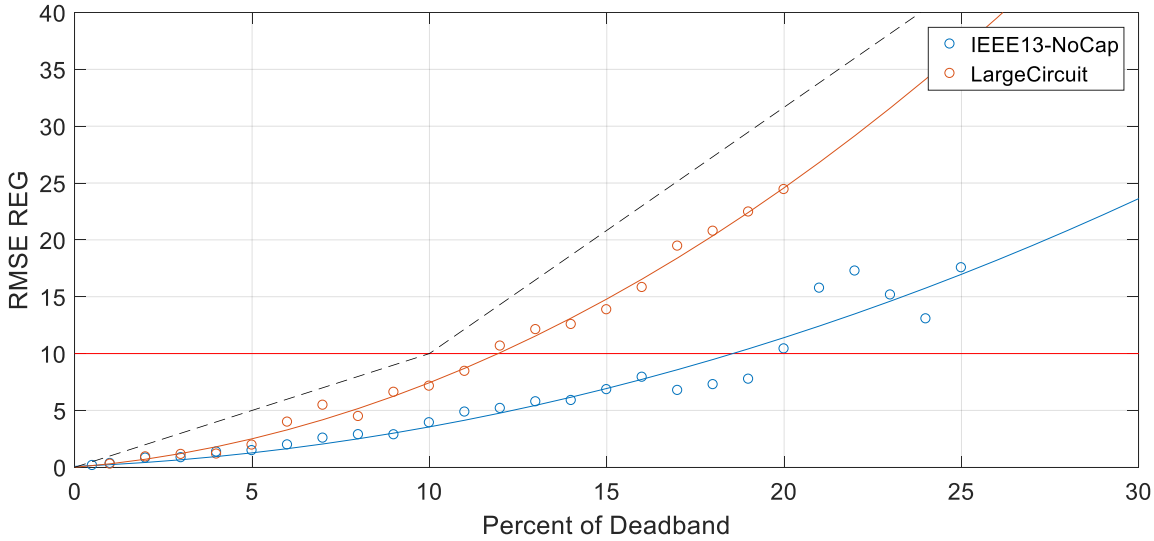


Figure 35 – RMS error of the number of controller action for all regulators versus the percent of deadband used to quantize the profiles for two additional test cases.

The test circuit where the capacitor controller is turned off yields accurate results compared to the results in Figure 34 although the total number of actions recorded is 8,373

for all regulators, which is lower than that of the IEEE13-1750 test case. This directly supports the claim that the error is more significant when there are complex interactions between controllers and not when there are fewer total number of actions. Furthermore, the error in the larger circuit is on the higher end of all simulated circuits even though the total number of actions is 20,312. Finally, these results agree with the observation that quantizing the profiles with a 10%-deadband strategy should yield error in the controller operations lower than 10%.

5.6 Discussion

5.6.1 Justification of the proposed VQ strategy

The results for the modified IEEE 13 bus test case show comparable performance from the kW-based and the proposed VQ strategy. Thus, one argument against the proposed method is that the kW-based approach is much simpler. Moreover, the results for the large distribution feeder test case show that the proposed strategy provides marginal improvements on the kW-based strategy. Although the proposed VQ strategy is more complex than the kW-based approach, the proposed sensitivity-based method does not impact the computational time of the algorithm since only $2N_{profile}$ power flows are required to compute the sensitivities which is less than 1 second. On the other hand, the advantage of this strategy is that the analysis is circuit specific and considers all the parameters of the simulation. One example would be two identical circuit with the same nominal load and PV output power but with different deadbands for their controllers. The kW-based would quantized the profiles identically for both circuit but the proposed VQ strategy would consider that difference. Moreover, the proposed VQ strategy also considers

the location of the power injection and controllers on the feeder. For instance, the two neighborhoods with distributed PV systems in the large distribution feeder test case have an identical aggregated output power (0.4 MW) but their impact on the voltage regulators (REG-a, -b, & -c) are different (see TABLE 9).

The work presented in this chapter discusses the effect of vector quantization on the accuracy of reported QSTS metrics. The proposed vector quantization strategy shows a practical method to quantize the different profiles based on sensitivity analysis of the controller input voltage signals. Because this sensitivity analysis is carried out prior to the time-series simulation, the profiles can be quantized according to their impact on various controllers. The advantage of this analysis is that it is embedded within the vector quantization algorithm discussed in Chapter 3 and does not require extensive knowledge on how the algorithm operates by the user. Ultimately, the work presented in this chapter makes the algorithm standalone so that it is implementable within commercial-grade distribution planning software such as OpenDSS or CYME. Since this research is part of a broader project collaboration with Sandia National Laboratories, both software companies have expressed interests in implementing this algorithm within their software.

5.6.2 Other applications of the sensitivity analysis

The sensitivity analysis discussed in this chapter can also be leveraged for other applications in distribution system modeling. For instance, another fast time-series algorithm reduces the computational time of QSTS simulations by representing the system with a linear sensitivity model [41], [51]. This collaborative work investigated the linear behavior of distribution systems when they operate near nominal voltage while still

including the effects of voltage-regulating devices. Similar to the vector quantization algorithm, each power flow solution is only impacted by the scaling multipliers of the power injections and the states of the voltage-regulating devices. Because these devices help maintain the voltage on the feeder near nominal, a linear sensitivity model can approximate the voltages and currents on the circuit without significant error. Furthermore, unlike a linear approximation of the power flow equations, this method considers the discontinuous nature of the system due to voltage-regulating equipment.

Another application of the sensitivity analysis and down-selection of the number of scenarios through vector quantization discussed in this chapter is in hosting capacity analysis [55]. Traditionally, hosting capacity analysis has been based on the peak/minimum load scenarios to capture the worst-case conditions. However, the probabilistic hosting capacity method proposed in [55] requires to run multiple loading scenarios over the range of load values. While the impact of the load may be different based on the specific feeder, the number of scenarios required for the probabilistic method will also differ and a similar approach to the sensitivity analysis discussed in this chapter is leveraged to adequately determine the number of loading scenarios needed. This work is outside the scope of this dissertation but more information can be found in [55].

CHAPTER 6. CONCLUSION AND FUTURE RESEARCH

6.1 Conclusion

Quasi-static time-series simulations can provide a detailed and realistic representation of the operation of distribution network with voltage-regulating equipment and solar PV systems. However, their computational burden prohibits their practical use by the industry. The work presented in this dissertation develops a vector quantization algorithm that drastically reduces the computational time of quasi-static time-series simulation.

The algorithm is capable of accurately modeling solar PV systems as well as the operation of voltage-regulating equipment (e.g. number of tap actions). By storing previously computed power flow solutions, the algorithm can reduce the number of times the non-linear 3 phase AC power flow equations are solved over the time horizon. Significant computational time reduction is attainable because of the cyclical behavior of the load and PV output power. A 99.5% computational time reduction or 200 times speed improvement is demonstrated on the IEEE 13 bus test case with an existing PV system with acceptable accuracies in the reported metrics. Namely, the error in the number of recorded actions by each controller remains below 5.7%, the voltage extremes below 0.0007 p.u. errors, the power loss below a 0.1% error, and the time outside ANSI limits below 5.0% errors. In addition to the attractive computational time reduction, the vector quantization algorithm provides real-time data compression for memory management by only storing unique power flow solutions compressing the data also by 99.5%.

Extensive effort is dedicated to address the scalability of the algorithm to model feeders of most complexities. First, the algorithm is not impacted by the size of the modeled distribution feeder since the computational time reduction is attained from reducing the number of computed power flows and not the time of an individual power flow solver. Second, the algorithm is capable of simulating feeder models with multiple load (i.e. residential, commercial, etc.) and PV profiles by preprocessing them as scenarios of power injections. Third, the capability of the algorithm is not limited by the number of controllable elements on the model through matrix decomposition. The scalability of the algorithm is demonstrated on a large realistic distribution system feeder with 2 load and 4 PV profiles, and 9 controllable voltage-regulating devices. Last, the algorithm is also capable of simulating feeders with different types of controllable devices that would be modeled with a QSTS simulation (e.g. smart inverter with VOLT/VAR capabilities).

The effect of vector quantization on the accuracy of the QSTS simulation is studied to propose a practical method to quantize the different profiles and reduce the computational time. The error in the power flow solution due to quantization is characterized for time-independent QSTS metrics such as voltage extremes or power losses. On the other hand, the discontinuous nature of the system created by voltage-regulating devices creates a challenge to estimate the error in time-dependent QSTS metrics such as number of controller actions or time outside ANSI. Thus, a strategy to quantize the different profiles is proposed based on a sensitivity analysis of the controller voltage input signals. Simulation results show that the proposed method is a practical approach to optimize speed and accuracy of QSTS simulations. Furthermore, the presented

vector quantization strategy is circuit-specific and does not require prior knowledge of the circuit or the algorithm to achieve attractive computational time reduction.

In summary, the vector quantization algorithm discussed in this dissertation can provide attractive computational time reduction and its robustness is discussed to model distribution feeders of most complexities. Moreover, a circuit-specific vector quantization method is proposed to make the algorithm standalone and implementable in commercial-grade distribution planning software. Thus, the vector quantization algorithm presented in this dissertation makes QSTS analysis a practical tool for utility planners and operators.

6.2 Contribution

The key contributions of the discussed work are the following:

- Reviewed the state of the art motivations and applications for QSTS simulation.
- Identified the challenges for fast detailed QSTS simulation and review the state of the art for fast QSTS methods.
- Developed a fast and robust QSTS simulation algorithm capable of modeling voltage-regulating devices and PV systems.
- Implemented the proposed fast time-series algorithm in MATLAB and validated its efficacy on a modified IEEE 13 bus test feeder.
- Addressed the scalability concerns of fast time-series algorithms and demonstrated the versatility of the proposed algorithm on a large realistic distribution feeder test case.
- Studied the effects of vector quantization on the accuracy of reported QSTS metrics.

- Developed a vector quantization strategy optimizing the speed and accuracy of the proposed algorithm.

For a list of publications, please refer to APPENDIX A.

6.3 Future work

Although the vector quantization algorithm discussed in this dissertation provides attractive computational speed, there are several research opportunities to further improve the performance of the algorithm. Namely, combining different fast time-series algorithms can provide additional computational time reduction. Currently, the algorithm steps through time and computes a new power flow solution once needed. However, multi-core computing could allow the algorithm to simultaneously compute similar power flow solutions that may later be experienced. By doing so, the algorithms could pre-compute some power flow solutions before they are experienced with any additional cores on the computer. Then, the computational time reduction of the algorithm would be at least the reduction presented in this dissertation.

In this dissertation, extensive work has been focused on demonstrating a practical and robust distribution system simulation tool. By making quasi-static time-series simulations practical, there is vast research opportunities for the application of this simulation tool. The application discussed in this dissertation focused on the impact analysis of new solar PV interconnections but other distributed energy resources could also be studied. For instance, wind turbines could be connected to medium-voltage distribution feeders and their impact should be studied prior to their connection. Furthermore, this fast time-series algorithm could also be used to determine appropriate settings in different

controllable devices on distribution feeders. Specifically, smart inverters with VOLT/VAR capabilities can be programmed differently based on their impact on a feeder. This fast time-series algorithm could facilitate that study.

APPENDIX A. PUBLICATIONS

Conference

J. Deboever, S. Grijalva, J. Peppanen, M. Rylander, and J. Smith, "Practical Data-Driven Methods to Improve the Accuracy and Detail of Hosting Capacity Analysis," in *IEEE Photovoltaic Specialists Conference (PVSC)2*, 2018.

J. Deboever, S. Grijalva, M. J. Reno, and R. J. Broderick, "Algorithms to Effectively Quantize Scenarios for PV Impact Analysis using QSTS Simulation," in *IEEE Photovoltaic Specialists Conference (PVSC)*, 2018, Submitted.

Q. Li, B. Mather, **J. Deboever**, and M. Reno, "Fast QSTS for Distributed PV Impact Studies using Vector Quantization and Variable Time-Steps," in *IEEE PES Innovative Smart Grid Technologies (ISGT) Conference (2018)*, Accepted.

J. Deboever, S. Grijalva, M. J. Reno, X. Zhang, and R. J. Broderick, "Scalability of the Vector Quantization Approach for Fast QSTS Simulation for PV Impact Studies," in *IEEE Photovoltaic Specialists Conference (PVSC)*, 2017.

X. Zhang, S. Grijalva, M. J. Reno, **J. Deboever**, and R. J. Broderick, "A Fast Quasi-Static Time Series (QSTS) Simulation Method for PV Impact Studies Using Voltage Sensitivities of Controllable Elements," in *IEEE Photovoltaic Specialists Conference (PVSC)*, 2017.

M. J. Reno, **J. Deboever**, and B. A. Mather, "Motivation and Requirements for Quasi-Static Time Series (QSTS) for Distribution System Analysis," in *IEEE PES General Meeting*, 2017.

Journal

J. Deboever, M. U. Qureshi, S. Grijalva, M. J. Reno, and R. J. Broderick, "A Sensitivity-based Quantization Method to Reduce Computational Time of QSTS simulations," *IEEE Trans. Sustain. Energy*, In preparation.

M. U. Qureshi, S. Grijalva, M. Reno, **J. Deboever**, X. Zhang, and R. Broderick, "A Fast Scalable Quasi-Static Time Series Analysis Method for PV Impact Studies using Linear Sensitivity Model," in *IEEE Transactions on Sustainable Energy*, Submitted.

J. Deboever, S. Grijalva, M. J. Reno, and R. J. Broderick, "Fast quasi-static time-series (QSTS) for yearlong PV impact studies using vector quantization," in *Solar Energy*, Accepted.

Technical Report

J. Deboever, X. Zhang, M. J. Reno, R. J. Broderick, and S. Grijalva, "Challenges in reducing computational time of QSTS simulations for distribution system analysis," SAND2017-5743, Albuquerque, NM, 2017.

Invited Speaker

J. Deboever, "Panel Session: Advances in Accelerated Distribution System Time-Series Analysis," in *IEEE PES General Meeting*, Chicago, IL, 2017.

Publications not relevant to the work

J. Deboever, J. Peppanen, A. Maitra, G. Damato, and J. Taylor, "Energy Storage as a Non-Wires Alternative for Deferring Distribution Capacity Investments," in *IEEE Transmission and Distribution Conference*, Accepted.

J. Deboever and S. Grijalva, "Energy Storage Dispatch under Different Ownership and Control Models," in *IEEE PES General Meeting*, 2017.

J. Deboever and S. Grijalva, "Optimal scheduling of large-scale price-maker energy storage," *2016 IEEE Power Energy Conf. Illinois, PECI 2016*, 2016.

J. Deboever, Santiago Grijalva, "Modeling and Optimal Scheduling of Integrated Thermal and Electrical Energy Microgrid," in *North American Power Symposium (NAPS 2016)*, 2016.

APPENDIX B. COMPUTATIONAL TIME ESTIMATION

The computational time reductions presented in this dissertation is estimated as if the algorithm is implemented within a commercial-grade distribution planning software (OpenDSS in this dissertation) and not in MATLAB or python through a COM object. The reason computational time reductions are presented as the percentage of the number of computed power flows instead of actual computational time is because there is a significant computational time associated with transferring information through the COM object. In this appendix, simulation results are presented supporting this justification.

For the vector quantization algorithm to be implemented in MATLAB and utilizing OpenDSS as the power flow engine, a number of information exchanges (hereafter referred as COM calls) occurs to access different measurements of the power flow solution. Each time a new power flow solution is computed, the following COM calls are needed to:

- set with the correct time-step in OpenDSS which has the specific scenario of power injections (load/PV); because the controller states change, the algorithm may come back to a specific scenario multiple times over the time horizon (single COM call),
- request OpenDSS to solve the power flow solution (single COM call),
- access the nodal voltage (single COM call),
- access the total line losses (single COM call),
- access the regulator control signals (3 COM calls per regulator) and,

- access the capacitor control signals (2 COM calls per capacitor).

Thus, for every power flow solved in the modified IEEE 13 test case, 15 COM calls are required to exchange information between OpenDSS and MATLAB. Furthermore, each time a controller takes an action, the states of the controller must be changed in the model in OpenDSS for it to impact the power flow solution. For each controller action taken by a regulator, a single COM call is needed and 2 COM calls are needed for any actions taken by a capacitor bank (capacitor name and status change).

A set of simulations are carried on the modified IEEE 13 bus test case where the load and PV profiles are uniformly quantized at 50, 55, 60, 65, …, 200 to support the claims above. First, a plot of the number of computed power flow solutions versus the number of COM calls demonstrates that there are 15 COM calls per solution and approximately 28,000 COM calls associated with the controllers changing states.

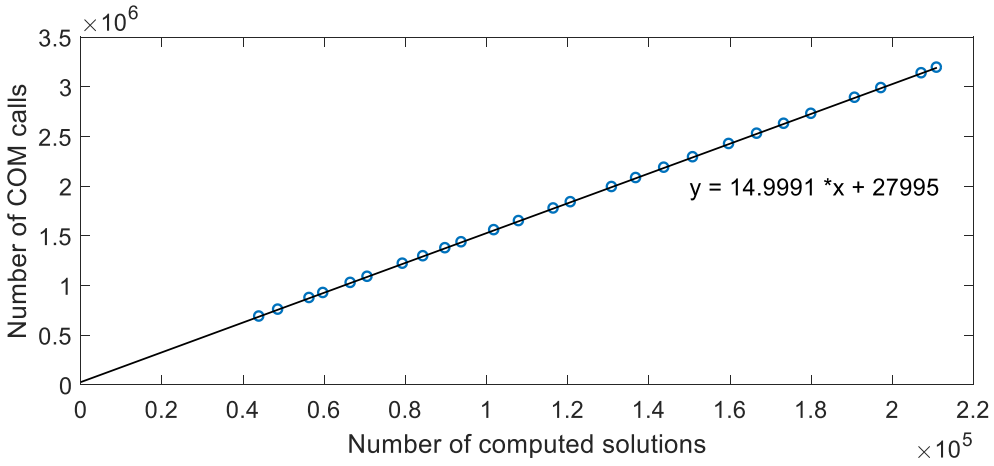


Figure 36 – Number of COM calls in terms of the number of computed solutions.

Second, a plot correlating the computational time to the number of COM calls to approximate the overhead computational time as if there were no COM calls (Figure 37). Based on the simulation results, the overhead computational time for the algorithm to go

through the logic that determines if a power flow solution already exists and through the logic of the various controllers is approximately 4.47 seconds.

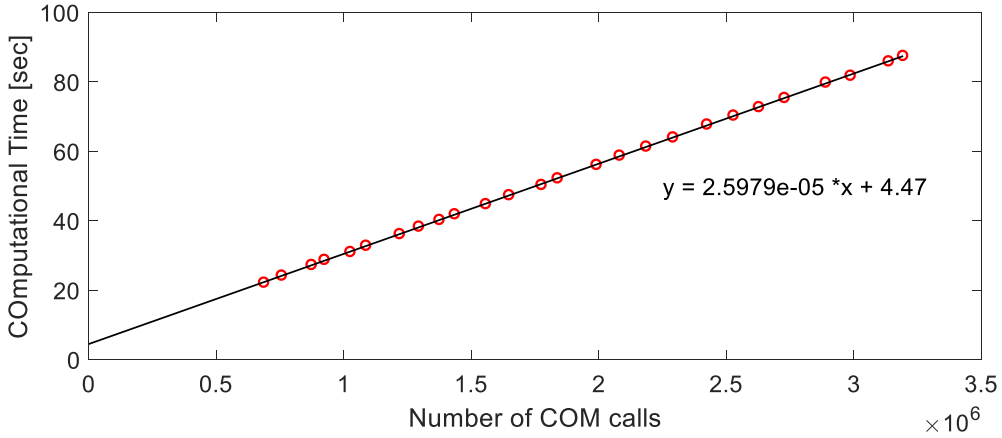


Figure 37 – Relationship between computational time and number of COM calls.

Alternatively, the overhead computational time can be estimated by re-computing the same QSTS simulation with the solution space already available. This avoids the power flow solver for every time-step over the time horizon. Since most COM calls occurs when solving a power flow solution, their impact on the computational time is significantly reduce and the overhead computational time of the algorithm can be estimated. The average computational time in this set of simulations is 4.40 seconds, which agrees with the estimation above.

APPENDIX C. BOX PLOT OF THE PV PROFILES

In this appendix, the hourly box plot for each of the four PV profiles in the realistic large distribution feeder test case are plotted. Each profile is different based on the location of the PV system it is associated to.

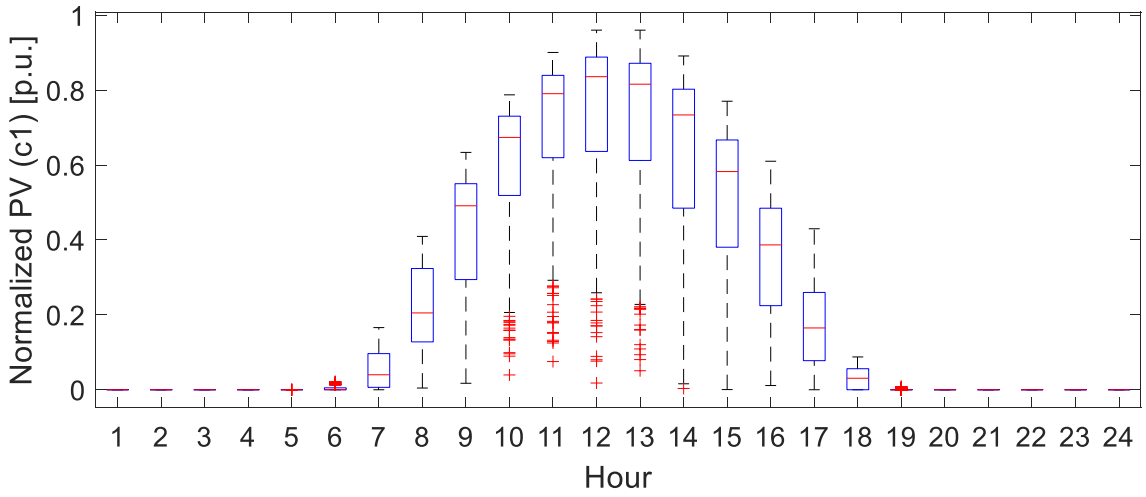


Figure 38 – Hourly box plot of the profile for the centralized PV system at location c1.

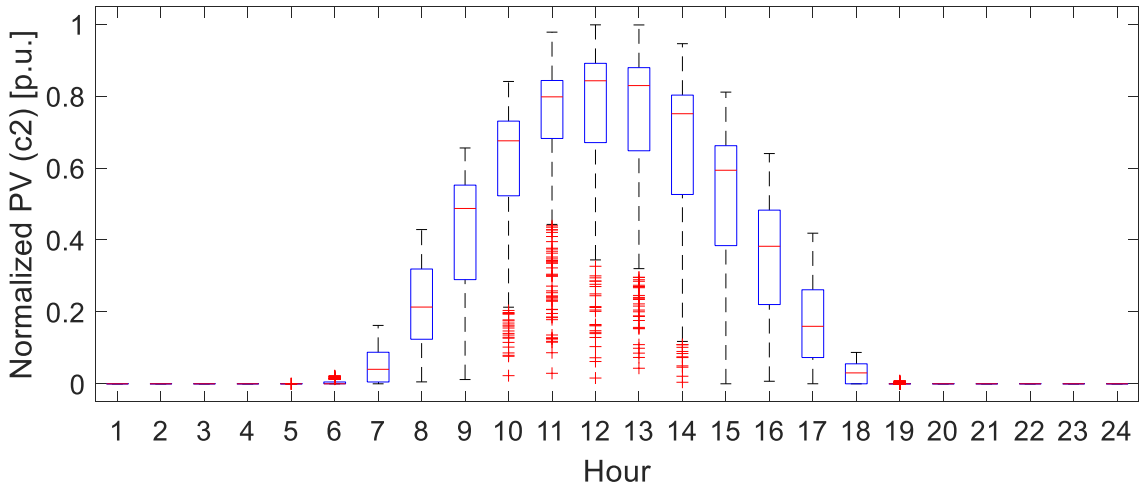


Figure 39 – Hourly box plot of the profile for the centralized PV system at location c1.

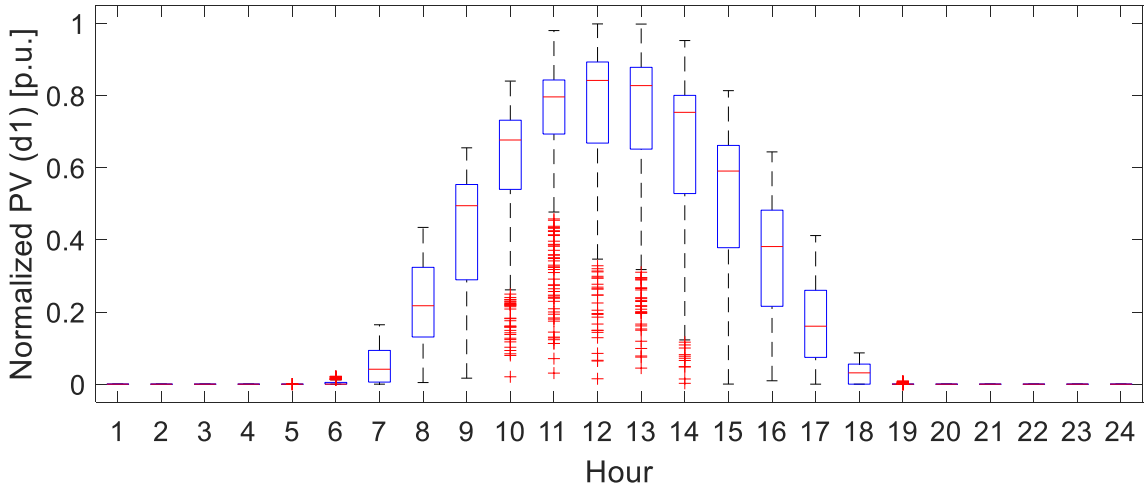


Figure 40 – Hourly box plot of the profile for the distributed PV systems at location d1.

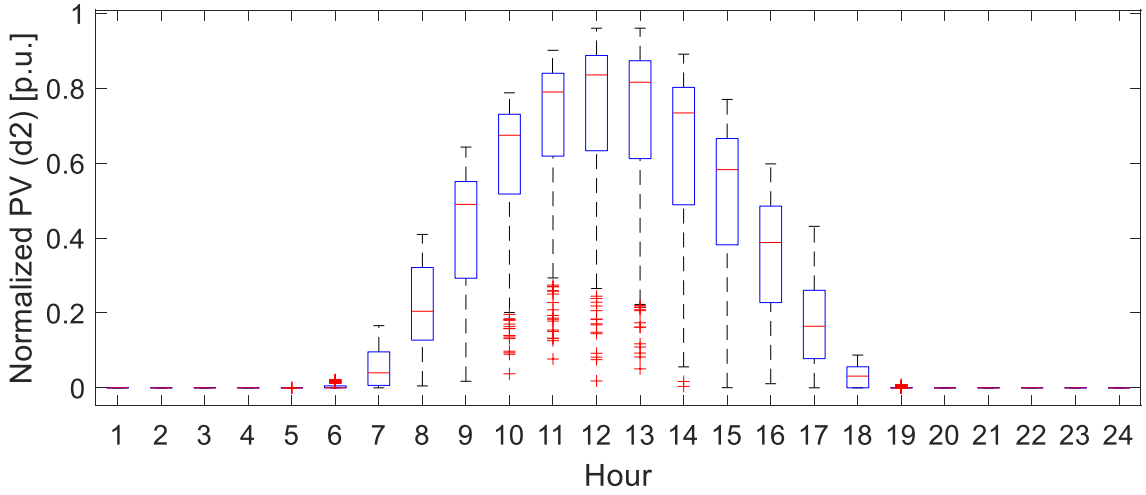


Figure 41 – Hourly box plot of the profile for the distributed PV systems at location d2.

REFERENCES

[1] U.S. Energy Information Administration, “Monthly Energy Review - November 2017,” 2017.

[2] S. Hanley, “Germany Breaks A Solar Record — Gets 85% Of Electricity From Renewables,” *Cleantechnica.com*, 2017. .

[3] DSIRE, “Residential Renewable Energy Tax Credit.” [Online]. Available: http://programs.dsireusa.org/system/program/detail/1235. [Accessed: 01-Jan-2017].

[4] Maui Electric, “Locational Value Map Tool,” 2017. [Online]. Available: https://www.mauielectric.com/clean-energy-hawaii/integration-tools-and-resources/locational-value-maps/oahu-locational-value-map-(lvm). [Accessed: 14-Mar-2017].

[5] IEEE P1547.7 D110, “Draft Guide to Conducting Distribution Impact Studies for Distributed Resource Interconnection,” 2013.

[6] M. A. Cohen, P. A. Kauzmann, and D. S. Callaway, “Effects of distributed PV generation on California’s distribution system, part 1: Engineering simulations,” *Sol. Energy*, vol. 128, pp. 139–152, 2016.

[7] M. J. Reno, J. Deboever, and B. A. Mather, “Motivation and Requirements for Quasi-Static Time Series (QSTS) for Distribution System Analysis,” in *IEEE PES General Meeting*, 2017.

[8] B. A. Mather, “Quasi-static time-series test feeder for PV integration analysis on distribution systems,” *IEEE Power Energy Soc. Gen. Meet.*, pp. 1–8, 2012.

[9] A. Pagnetti and G. Delille, “A simple and efficient method for fast analysis of renewable generation connection to active distribution networks,” *Electr. Power Syst. Res.*, vol. 125, pp. 133–140, 2015.

[10] R. Yan, B. Marais, and T. K. Saha, “Impacts of residential photovoltaic power fluctuation on on-load tap changer operation and a solution using DSTATCOM,” *Electr. Power Syst. Res.*, vol. 111, pp. 185–193, 2014.

[11] M. J. E. Alam, K. M. Muttaqi, and D. Sutanto, “An approach for online assessment of rooftop solar PV impacts on low-voltage distribution networks,” *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 663–672, 2014.

[12] M. Baggu, R. Ayyanar, and D. Narang, “Feeder model validation and simulation for high-penetration photovoltaic deployment in the Arizona Public Service system,” in *IEEE Photovoltaic Specialists Conference*, 2014, pp. 2088–2093.

[13] R. J. Broderick, J. E. Quiroz, M. J. Reno, Abraham Ellis, J. Smith, and R. Dugan, “Time Series Power Flow Analysis for Distribution Connected PV Generation,” SAND2013-0537, Albuquerque, NM, 2013.

[14] D. Paradis, F. Katiraei, and B. Mather, “Comparative analysis of time-series studies and transient simulations for impact assessment of PV integration on reduced IEEE 8500 node feeder,” *IEEE Power Energy Soc. Gen. Meet.*, pp. 1–5, 2013.

[15] J. E. Quiroz, M. J. Reno, and R. J. Broderick, “Time series simulation of voltage regulation device control modes,” in *IEEE Photovoltaic Specialists Conference*, 2013, pp. 1700–1705.

[16] J. W. Smith, R. Dugan, and W. Sunderman, “Distribution modeling and analysis of high penetration PV,” *2011 IEEE Power Energy Soc. Gen. Meet.*, pp. 1–7, 2011.

[17] J. W. Smith, W. Sunderman, R. Dugan, and B. Seal, “Smart inverter volt/var control functions for high penetration of PV on distribution systems,” *2011 IEEE/PES Power Syst. Conf. Expo. PSCE 2011*, pp. 1–6, 2011.

[18] A. Hariri, M. O. Faruque, R. Soman, and R. Meeker, “Impacts and interactions of voltage regulators on distribution networks with high PV penetration,” *2015 North Am. Power Symp. NAPS 2015*, 2015.

[19] A. Khoshkbar-Sadigh and K. M. Smedley, “The necessity of time-series simulation for investigation of large-scale solar energy penetration,” *IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf.*, pp. 1–5, 2015.

[20] T. Boehme, A. R. Wallace, and G. P. Harrison, “Applying Time Series to Power Flow Analysis in Networks With High Wind Penetration,” *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 951–957, 2007.

[21] M. A. Abdel-warth, M. Abdel-akher, and M. M. Aly, “Quasi-Static Time-Series Simulation of Congested Power Systems with Wind Power Plant,” in *17th International Middle East Power System Conference MEPCON’15*, 2015.

[22] J. R. Aguero, P. Chongfuangprinya, S. Shao, L. Xu, F. Jahanbakhsh, and H. L. Willis, “Integration of Plug-in Electric Vehicles and distributed energy resources on power distribution systems,” *2012 IEEE Int. Electr. Veh. Conf.*, pp. 1–7, 2012.

[23] S. Shao, F. Jahanbakhsh, J. R. Aguero, and L. Xu, “Integration of PEVs and PV-DG in power distribution systems using distributed energy storage—Dynamic analyses,” in *IEEE PES Innovative Smart Grid Technologies (ISGT)*, 2013.

[24] M. Kleinberg, J. Harrison, and N. Mirhosseini, “Using energy storage to mitigate PV impacts on distribution feeders,” in *IEEE PES Innovative Smart Grid Technologies (ISGT)*, 2014.

[25] Meghasai, S. Monger, R. Vega, and H. Krishnaswami, “Simulation of smart functionalities of photovoltaic inverters by interfacing OpenDSS and Matlab,” *2015 IEEE 16th Work. Control Model. Power Electron. COMPEL 2015*, no. Lv, 2015.

[26] R. Michael, “Thesis: Online optimization of capacitor switching in electric power distribution systems,” Drexel University, 2015.

[27] M. Kraiczy, M. Braun, G. Wirth, S. Schmidt, and J. Brantl, “Interferences between local voltage control strategies of a hv/mv- transformer and distributed generators,” in *Photovoltaic Solar Energy Conference and Exhibition*, 2013, vol. 4, no. 1, pp. 1–15.

[28] K. P. Schneider, J. C. Fuller, and D. Chassin, “Evaluating conservation voltage reduction: An application of GridLAB-D: An open source software package,” *IEEE Power Energy Soc. Gen. Meet.*, pp. 1–6, 2011.

[29] Y. P. Agalgaonkar, B. C. Pal, and R. A. Jabr, “Distribution voltage control considering the impact of PV generation on tap changers and autonomous regulators,” *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 182–192, 2014.

[30] J. Deboever, X. Zhang, M. J. Reno, R. J. Broderick, and S. Grijalva, “Challenges in reducing computational time of QSTS simulations for distribution system analysis,” SAND2017-5743, Albuquerque, NM, 2017.

[31] W. H. Kersting, *Distribution system modeling and analysis*, 3rd ed. Las Cruces, NM: CRC Press, 2012.

[32] J. B. Ward and H. W. Hale, “Digital Computer Solution of Power-Flow Problems,” *AIEE Trans. (Power Appl. Syst.)*, vol. 75, pp. 398–404, 1956.

[33] R. C. Dugan, “OpenDSS Manual,” 2016.

[34] M. E. Baran and F. F. Wu, “Network reconfiguration in distribution systems for loss reduction and load balancing,” *IEEE Trans. Power Deliv.*, vol. 4, no. 2, pp. 1401–1407, 1989.

[35] F. Therrien, M. Belletête, J. Lacroix, and M. J. Reno, “Algorithmic Aspects of a Commercial-Grade Distribution System Load Flow Engine,” in *IEEE Photovoltaic Specialists Conference (PVSC)*, 2017.

[36] D. Montenegro, G. A. Ramos, and S. Bacha, “A-Diakoptics for the Multicore Sequential-Time Simulation of Microgrids Within Large Distribution Systems,” *IEEE Trans. Smart Grid*, pp. 1–9, 2015.

[37] C. D. López, B. Idlbi, T. Stetz, and M. Braun, “Shortening Quasi-Static Time-Series Simulations for Cost-Benefit Analysis of Low Voltage Network Operation with Photovoltaic Feed-In,” in *Power and Energy Student Summit (PESS)*, 2015.

[38] Z. K. Pecenak, V. R. Disfani, M. J. Reno, and J. Kleissl, “Multiphase Distribution Feeder Reduction,” *IEEE Trans. Power Syst.*, pp. 1–1, 2017.

[39] M. J. Reno, K. Coogan, R. Broderick, and S. Grijalva, “Reduction of distribution feeders for simplified PV impact studies,” in *IEEE Photovoltaic Specialists Conference*, 2013, pp. 2337–2342.

[40] American National Standards Institute (ANSI), “American National Standard for Electric Power Systems and Equipment - Voltage Ratings (60Hz), ANSI C84.1-2006.” 2006.

[41] X. Zhang, S. Grijalva, M. J. Reno, J. Deboever, and R. J. Broderick, “A Fast Quasi-Static Time Series (QSTS) Simulation Method for PV Impact Studies Using Voltage Sensitivities of Controllable Elements,” in *IEEE Photovoltaic Specialists Conference (PVSC)*, 2017.

[42] M. J. Reno and K. Coogan, “Grid Integrated Distributed PV (GridPV),” SAND2013-6733, 2013.

[43] National Renewable Energy Laboratory, “Global Horizontal Irradiance data from NREL Oahu.” [Online]. Available: http://www.nrel.gov/midc/oahu_archive/.

[44] M. Lave, W. Hayes, A. Pohl, and C. W. Hansen, “Evaluation of global horizontal

irradiance to plane-of-array irradiance models at locations across the United States," *IEEE J. Photovoltaics*, vol. 5, no. 2, pp. 597–606, 2015.

[45] J. Deboever, S. Grijalva, M. J. Reno, and R. J. Broderick, "Fast quasi-static time-series (QSTS) for yearlong PV impact studies using vector quantization," *Sol. Energy*, 2018.

[46] M. J. Reno, M. Lave, J. E. Quiroz, and R. J. Broderick, "PV Ramp Rate Smoothing Using Energy Storage to Mititgate Increased Voltage regulator Tapping," *IEEE Photovolt. Spec. Conf.*, 2016.

[47] J. Seuss, M. J. Reno, R. J. Broderick, and S. Grijalva, "Analysis of PV Advanced Inverter Functions and Setpoints under Time Series Simulation," SANDSAND2016-4856, Albuquerque, NM, 2016.

[48] M. Lave, J. Kleissl, and J. S. Stein, "A wavelet-based variability model (WVM) for solar PV power plants," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 501–509, 2013.

[49] A. Nagarajan, B. Palmintier, F. Ding, B. Mather, and M. Baggu, "Improving advanced inverter control convergence in distribution power flow," *NAPS 2016 - 48th North Am. Power Symp. Proc.*, 2016.

[50] B. A. Mather, "Fast Determination of Distribution-Connected PV Impacts Using a Variable Time-Step Quasi-Static Time-Series Approach," in *IEEE Photovoltaic Specialists Conference (PVSC)*, 2017.

[51] M. U. Qureshi, S. Grijalva, M. J. Reno, J. Deboever, and R. J. Broderick, “A Fast Scalable Quasi-Static Time Series Analysis Method for PV Impact Studies using Linear Sensitivity Model,” *IEEE Trans. Sustain. Energy2*.

[52] Q. Li, B. A. Mather, J. Deboever, and M. J. Reno, “Fast QSTS for Distributed PV Impact Studies using Vector Quantization and Variable Time-Steps,” in *IEEE PES Innovative Smart Grid Technologies (ISGT)2*, 2018.

[53] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Springer US, 1992.

[54] A. C. Bovik, “Chapter 8.1: Image Quantization, Halftoning, and Printing,” in *Handbook of Image and Video Processing*, 2nd ed., Academic Press, 2010, p. 926.

[55] J. Deboever, S. Grijalva, J. Peppanen, M. Rylander, and J. Smith, “Practical Data-Driven Methods to Improve the Accuracy and Detail of Hosting Capacity Analysis,” in *IEEE Photovoltaic Specialists Conference (PVSC)2,* 2018.

[56] J. Deboever, M. U. Qureshi, S. Grijalva, M. J. Reno, and R. J. Broderick, “A Sensitivity-based Quantization Method to Reduce Computational Time of QSTS simulations,” *IEEE Trans. Sustain. Energy*.

[57] J. Deboever, S. Grijalva, M. J. Reno, X. Zhang, and R. J. Broderick, “Scalability of the Vector Quantization Approach for Fast QSTS Simulation for PV Impact Studies,” in *IEEE Photovoltaic Specialists Conference (PVSC)*, 2017.

[58] J. Deboever, S. Grijalva, M. J. Reno, and R. J. Broderick, “Algorithms to Effectively

Quantize Scenarios for PV Impact Analysis using QSTS Simulation," in *IEEE Photovoltaic Specialists Conference (PVSC),* 2018.

[59] J. Deboever, "Panel Session: Advances in Accelerated Distribution System Time-Series Analysis," in *IEEE PES General Meeting (2017)*, 2017.

[60] J. Deboever and S. Grijalva, "Modeling and Optimal Scheduling of Integrated Thermal and Electrical Energy Microgrid," *North Am. Power Sysmposium*, 2016.

[61] J. Deboever and S. Grijalva, "Optimal scheduling of large-scale price-maker energy storage," *2016 IEEE Power Energy Conf. Illinois, PECI 2016,* 2016.

[62] J. Deboever and S. Grijalva, "Energy Storage Dispatch under Different Ownership and Control Models," in *IEEE PES General Meeting*, 2017.

[63] J. Deboever, J. Peppanen, A. Maitra, G. Damato, and J. Taylor, "Energy Storage as a Non-Wires Alternative for Deferring Distribution Capacity Investments," in *IEEE Transmission and Distribution Conference*, 2018.